

**WEST****Create A Case**

---

Select?	Database	Query	Plural	Op	Thesaurus	Set Name
<input checked="" type="checkbox"/>	USPT	anti-tumor antibody	YES	ADJ	ASSIGNEE	L1
<input checked="" type="checkbox"/>	USPT	anti-tumorigenic antibody	YES	ADJ	ASSIGNEE	L2
<input checked="" type="checkbox"/>	USPT	antibody and anti-tumorigenic activity	YES	ADJ	ASSIGNEE	L3

Please enter the case name:

---

---

**Rules for naming Cases**

- Case names can only contain alphanumeric characters including underscore (\_).
- Any other special characters or punctuation characters will be automatically removed prior to saving the case.
- All white space characters will be replaced by an underscore.

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2003, 16:08:55 ; Search time 47.5 Seconds  
(without alignments)

53.300 Million cell updates/sec

Title: US-09-824-647-3

Perfect score: 96  
Sequence: 1 KKVIAPIRRLLPQILKSDT 19

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96	100.0	19	20 AAW85478	Mouse GP88 autocri
2	96	100.0	589	12 AAR14327	Mouse epithelin pr
3	89	92.7	589	20 AAW85474	Mouse GP88 autocri
4	89	92.7	589	23 AAE20520	Mouse granulin/epi
5	62	64.6	589	12 AAR14325	Rat epithelin prec
6	47	49.0	650	22 ABG07964	Novel human diagn
7	46	47.9	824	22 AAE20251	Human ADAM8. Homo
8	46	47.9	824	23 AA086156	Human PRO1686 poly
9	44	45.8	19	20 AAW85480	Human GP88 autocri
10	44	45.8	29	18 AAW18298	Human calpastatin

11	44	45.8	413	22 AAB94550	Human protein sequ
12	44	45.8	463	18 AAW19394	Human calpastatin
13	44	45.8	593	12 AAR14326	Human epithelin pr
14	44	45.8	593	14 AAR48673	Granulin sequence.
15	44	45.8	593	20 AAW85475	Human GP88 autocri
16	44	45.8	593	23 AAE20521	Human granulin/epi
17	44	45.8	621	21 AAB43971	Human cancer asoc
18	43	44.8	108	21 AAG16784	Arabidopsis thalia
19	43	44.8	131	21 AAG16783	Arabidopsis thalia
20	43	44.8	142	21 AAG16782	Arabidopsis thalia
21	43	44.8	148	21 AAG55018	Arabidopsis thalia
22	43	44.8	154	22 AAE58583	Propionibacterium
23	43	44.8	595	23 AAE16337	Human POLY1 protei
24	43	44.8	595	23 AAE16338	Human POLY2 protei
25	43	44.8	595	23 AAE16339	Human POLY3 protei
26	43	44.8	596	22 AAG67544	Amino acid sequenc
27	43	44.8	596	23 AAE22723	Human cardiac myos
28	43	44.8	596	23 AAE22847	Human cardiac myos
29	43	44.8	596	23 AAE22848	Human cardiac myos
30	43	44.8	596	23 AAE22849	Human cardiac myos
31	43	44.8	596	23 AAE22850	Human cardiac myos
32	43	44.8	596	23 AAE16260	Human kinase PKIN-
33	43	44.8	596	23 AAE16340	Human POLY4 protei
34	43	44.8	612	22 AA003521	Human protein kina
35	43	44.8	692	22 ABB69139	Drosophila melanog
36	42	43.8	57	22 AAU58470	Propionibacterium
37	42	43.8	97	22 ABG24494	Novel human diagn
38	42	43.8	97	23 ABP47810	SRV-2 protein rela
39	42	43.8	255	22 ABG15125	Novel human diagn
40	42	43.8	298	22 ABB66015	Drosophila melanog
41	42	43.8	321	22 ABG20636	Novel human diagn
42	42	43.8	416	22 AAG67388	Amino acid sequenc
43	42	43.8	416	22 AAE02177	Mycobacterium tube
44	42	43.8	440	23 ABP47767	Protein #13 relate
45	42	43.8	501	22 ABG15126	Novel human diagn

#### ALIGNMENTS

RESULT 1

AAW85478  
ID AAW85478 standard; Peptide: 19 AA.

XX AC AAW85478;

XX DT 15-MAR-1999 (first entry)

XX DE Mouse GP88 autocrine growth factor antigenic peptide K19T.

XX DE GP88; granulin; epithelin; mouse; growth factor; autocrine; tumour;  
KW cancer; viral infection; antagonist; therapy; diagnosis; antigen;  
KW antibody.

XX OS Mus sp.

XX PN WO9852607-A1.

XX PD 26-NOV-1998.

XX PF 22-MAY-1998; 98WO-US10555.

XX PR 16-DEC-1997; 97US-0991862.

XX PR 23-MAY-1997; 97US-0863079.

XX (SERR/) SERRERO G.

XX Serrero G;

XX WPI; 1999-045276/04.

XX Composition containing antagonist of growth factor GP88 - useful for  
PT treating cancer and viral diseases and also for diagnosing disease  
PT

from altered GP88 expression

Example 8; Page 45; 86pp; English.

This is the amino acid sequence of peptide K19T, comprising amino acid residues K344-T362 of murine GP88 (see AAW85474). GP88 is an 88 kDa glycoprotein autocrine growth factor that is expressed in a tightly regulated manner in normal cells, is overexpressed and unregulated in highly tumorigenic cells derived from normal cells, and which acts as a stringently required growth stimulator for the tumorigenic cells. K19T was used to raise neutralising antibodies to GP88. Antagonists to GP88, such as anti-GP88 antibodies, are used to treat diseases associated with increased expression of GP88, particularly cancer but also viral infections. Anti-GP88 antibodies can also be used as diagnostic reagents and to deliver toxins or other compounds to GP88-expressing cells.

XX Sequence 19 AA;

Query Match 100.0%; Score 96; DB 20; Length 19;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKVIAPRRLLDPQILKSDT 19  
 DB 1 KKVIAPRRLLDPQILKSDT 19

RESULT 2

ID AAR14327  
 ID AAR14327 standard; Protein; 589 AA.

AC AAR14327;

DT 17-JAN-1992 (first entry)

DE Mouse epithelin precursor.

KW ET; growth regulation; inhibition; stimulation.

OS Mus musculus.

FH Key Location/Qualifiers

FT Protein 1..589

FT /label= precursor

FT /note= "claim 21, page 55"

FT Protein 280..335

FT /label= EP-1

FT /note= "claim 22, page 55"

FT Protein 205..261

FT /label= EP-2

FT /note= "claim 23, page 55"

FT Peptide 59..114

FT /label= EP

FT /note= "claim 24, page 55"

FT Peptide 123..179

FT /label= EP

FT /note= "claim 25, page 55"

FT Peptide 362..416

FT /label= EP

FT /note= "claim 26, page 56"

FT Peptide 440..495

FT /label= EP

FT /note= "claim 27, page 56"

FT Peptide 515..570

FT /label= EP

FT /note= "claim 28, page 56"

PN W09115510-A.

XX 17-OCT-1991.

XX 03-APR-1991; 91WO-US02321.

XX 13-MAR-1991; 91US-0083796.  
 PR 03-APR-1990; 90US-0504508.

XX (BRIM ) BRISTOL-MYERS SQUIB.

XX Shoyab M, Plowman GD;

XX WPI; 1991-325168/44.

XX N-PSDB; AAQ14340.

XX New cysteine-rich growth modulating proteins, epithelins - useful as inhibitors of neoplastic cell growth and to promote wound healing and treat psoriasis

XX Disclosure; Fig 23; 97pp; English.

XX ET-1 is a bifunctional growth regulator, capable of stimulating the growth of some cell types while inhibiting the growth of others. ET-2 is functionally similar to ET-1 w.r.t. growth inhibitory bioactivity. In contrast, however, ET-2 is apparently not capable of eliciting the growth stimulatory activity characteristic of ET-1 and, in fact, antagonises this ET-1 activity.

XX See also AAQ14338-40, AAQ14952-53, AAR14328-9 and AAR15315-20.

XX Sequence 589 AA;

Query Match 100.0%; Score 96; DB 12; Length 589;

Best Local Similarity 100.0%; Pred. No. 1.4e-06;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKVIAPRRLLDPQILKSDT 19

DB 344 KKVIAPRRLLDPQILKSDT 362

RESULT 3

AAW85474

ID AAW85474 standard; Protein; 589 AA.

XX AC AAW85474;

XX DT 15-MAR-1999 (first entry)

XX Mouse GP88 autocrine growth factor.

DE GP88; granulin; epithelin; mouse; growth factor; autocrine; tumour;

KW cancer; viral infection; antagonist; therapy; diagnosis.

XX Mus sp.

XX Key Location/Qualifiers

FT Misc-difference 8 /note= "encoded by CTG"

FT Misc-difference 54 /note= "encoded by AGC"

FT Misc-difference 377 /note= "encoded by TGA"

FT Peptide 208..219

FT /note= "P12T peptide used to raise antibody"

FT Peptide 344..362

FT /note= "K19T peptide, used to raise antibody"

FT Peptide 562..575

FT /note= "S14R peptide, used to raise antibody"

XX W09852607-A1.

XX 26-NOV-1998.

XX 22-MAY-1998; 98WO-US10555.

XX 16-DEC-1997; 97US-0991862.

XX 23-MAY-1997; 97US-0863079.

XX PA (SERR/) SERRERO G.  
 XX PI Serrero G;  
 XX PR WPI: 1999-045276/04.  
 XX DR N-PSDB; AAV82824.  
 XX PA Composition containing antagonist of growth factor GP88 - useful for  
 PT treating cancer and viral diseases and also for diagnosing disease  
 PT from altered GP88 expression  
 XX PS Example 5; Fig 8A-D; 86pp; English.  
 XX CC This is the amino acid sequence of murine GP88, an 88 kDa  
 CC glycoprotein autocrine growth factor and epithelin/granulin  
 CC precursor that is expressed in a tightly regulated manner in normal  
 CC cells. It is overexpressed and unregulated in highly tumorigenic cells  
 CC derived from normal cells, and which acts as a stringently required  
 CC growth stimulator for the tumorigenic cells. Inhibition of GP88  
 CC expression or action in the tumorigenic cells results in an  
 CC inhibition of the tumorigenic properties of the overproducing  
 CC cells. Murine GP88 cDNA (see AAV82824) was obtained from the highly  
 CC tumorigenic PC cell line. Antagonists to GP88 are used to treat  
 CC diseases associated with increased expression of GP88, particularly  
 CC cancer but also viral infections. Fragments of GP88 are used to  
 CC raise specific antibodies (used as antagonists, as diagnostic  
 CC reagents and for delivering toxins or other compounds to GP88-  
 CC expressing cells) and to screen for antibodies. Methods are  
 CC provided for diagnosing disease, or determining susceptibility to  
 CC disease, resulting from altered GP88 activity.  
 XX CC Sequence 589 AA;  
 XX SQ Query Match 92.7%; Score 89; DB 20; Length 589;  
 XX Best Local Similarity 94.7%; Pred. No. 1.7e-05;  
 XX Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 OY 1 KKVIAPRLPPQILKSDT 19  
 DB 344 KKVIAPRLPPQILKSDT 362  
 RESULT 4  
 AAE20520  
 ID AAE20520 standard; Protein; 589 AA.  
 XX AC AAE20520;  
 XX DT 01-JUL-2002 (first entry)  
 XX DE Mouse granulin/epithelin precursor (GP88) protein.  
 XX KW Mouse; granulin precursor; GP88; cytostatic; tumorigenicity; tamoxifen;  
 XX KW antineoplastic; antioestrogen therapy; skin cancer.  
 XX OS Mus sp.  
 XX FH Key Location/Qualifiers  
 XX FT Misc-difference 335 /label= Unknown  
 XX FT /note= "Encoded by ATG"  
 XX FT Region 344..362  
 XX FT /note= "Regions used as antigens to raise anti-GP88  
 XX FT antibodies"  
 XX FT Region 562..575  
 XX FT /note= "Regions used as antigens to raise anti-GP88  
 XX FT antibodies"  
 XX FT Misc-difference 586  
 XX FT /note= "Encoded by ACA"  
 XX PN US2002025543-A1.  
 XX FT

28-FEB-2002.  
 15-JUN-2001; 2001US-0880842.  
 23-MAY-1997; 97US-0863079.  
 08-DEC-1999; 99US-0456886.  
 (SERR/) SERRERO G.  
 Serrero G;  
 WPI: 2002-267529/31.  
 N-PSDB; AAD32849.  
 Diagnosing tumorigenicity in a human, comprising obtaining a cell  
 sample, detecting GP88 in the cells, and determining the number of GP88  
 positive cells in the sample -  
 Disclosure; Fig 8; 50pp; English.  
 The invention relates to a method for diagnosing tumorigenicity in a  
 human. The method comprises obtaining a biological sample containing  
 cells from the patient, detecting GP88 in the cells of the sample,  
 and determining the number of GP88 positive cells in the sample, and  
 determining the ratio of GP88 positive cells to the total number of cells  
 in the sample. The invention also relates to a method for determining  
 if a human patient is resistant to the antineoplastic effects of  
 antioestrogen therapy. The method is useful for diagnosing  
 tumorigenicity in a sample, such as blood, serum, plasma, urine, nipple  
 aspirate, cerebrospinal fluid, liver, kidney, breast, bone, brain, lung,  
 colon, or skin cancer. The method can be used to treat or prevent  
 re-occurrence of cancer in a patient, by administering tamoxifen if the  
 sample contains less than 10 % GP88, or less than 5 % GP88 positive  
 cells. The present sequence is mouse granulin/epithelin precursor (GP88)  
 protein.  
 Query Match 92.7%; Score 89; DB 23; Length 589;  
 Best Local Similarity 94.7%; Pred. No. 1.7e-05;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 OY 1 KKVIAPRLPPQILKSDT 19  
 DB 344 KKVIAPRLPPQILKSDT 362  
 RESULT 5  
 AAR14325  
 ID AAR14325 standard; Protein; 589 AA.  
 XX AC AAR14325;  
 XX DT 17-JAN-1992 (first entry)  
 XX DE Rat epithelin precursor.  
 XX KW ET; growth regulation; inhibition; stimulation.  
 XX OS Rattus rattus.  
 XX FH Key Location/Qualifiers  
 XX FT Protein 1..589  
 XX FT /label= precursor  
 XX FT /note= "Claim 11, page 54"  
 XX FT Protein 280..335  
 XX FT /label= EP-1  
 XX FT /note= "claim 12, page 54"  
 XX FT Protein 205..261  
 XX FT /label= EP-2  
 XX FT /note= "claim 13, page 54"  
 XX FT Peptide 59..114  
 XX FT /label= EP



FT PT Peptide /note= "claim 14, page 54"  
123..179  
/label= EP  
FT PT Peptide /note= "claim 15, page 54"  
362..416  
/label= EP  
FT PT Peptide /note= "claim 16, page 54"  
440..495  
/label= EP  
FT PT Peptide /note= "claim 17, page 54"  
515..570  
/label= EP  
FT PT Peptide /note= "claim 18, page 55"  
WO9115510-A.  
XX PN 17-OCT-1991.  
XX PD 03-APR-1991; 91WO-US02321.  
XX PF 13-MAR-1991; 91US-0083796.  
XX PR 03-APR-1990; 90US-0504508.  
XX PA (BRIM ) BRISTOL-MYERS SQUIB.  
XX PI Shoyab M, Plowman GD;  
XX XX WPI; 1991-325168/44.  
DR N-PSDB; AAQ14338.  
XX XX New cysteine-rich growth modulating proteins, epithelins - useful  
PT as inhibitors of neoplastic cell growth and to promote wound  
PT healing and treat psoriasis  
XX XX Disclosure; Fig 18; 97pp; English.  
XX CC ET-1 is a bifunctional growth regulator, capable of stimulating  
CC the growth of some cell types while inhibiting the growth of others.  
CC ET-2 is functionally similar to ET-1 w.r.t. growth inhibitory  
CC bioactivity. In contrast, however, ET-2 is apparently not capable of  
CC eliciting the growth stimulatory activity characteristic of ET-1 and,  
CC in fact, antagonises this ET-1 activity.  
CC See also AAQ14338-40, AAQ14952-53, AAR14328-9 and AAR15315-20.  
XX XX  
SQ Sequence 589 AA;  
Query Match 64.6%; Score 62; DB 12; Length 589;  
Best Local Similarity 72.2%; Pred. No. 0.26;  
Matches 13; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
QY 1 KKVIAPRRLPDQPILKSD 18  
DB 344 KKVTAASLSLPDQPILKND 361  
RESULT 6  
ABG07964  
ID ABG07964 standard; Protein; 650 AA.  
XX AC ABG07964;  
XX DT 13-FEB-2002 (first entry)  
XX DE Novel human diagnostic protein #7955.  
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX OS Homo sapiens.  
XX XX WO200175067-A2.  
PN 11-OCT-2001.  
XX PD

XX 30-MAR-2001; 2001WO-US08631.  
XX 31-MAR-2000; 2000US-0540217.  
PR 23-AUG-2000; 2000US-0649167.  
XX (HYSE-) HYSEQ INC.  
XX Drmanac RT.; Liu C, Tang YT;  
XX WPI; 2001-639362/73.  
DR N-PSDB; AAS72151.  
XX PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity  
XX PS Claim 20; SEQ ID No 38323; 103pp; English.  
XX CC The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human  
CC amino acid sequences. ABG00010-ABG30377 represent novel human  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX SQ Sequence 650 AA;  
Query Match 49.0%; Score 47; DB 22; Length 650;  
Best Local Similarity 66.7%; Pred. No. 60;  
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 4 IAPRLPDQPIL 15  
DB 291 LAPRLPDQPVL 302  
RESULT 7  
AAB20251  
ID AAB20251 standard; Protein; 824 AA.  
XX AC AAB20251;  
XX DT 14-MAY-2001 (first entry)  
XX DE Human ADAM8.  
XX KW ADAM8; human; tumour; cancer; therapy; diagnosis; antitumour.  
XX OS Homo sapiens.  
XX PN WO200109189-A2.  
XX XX 08-FEB-2001.  
XX PF 27-JUL-2000; 2000WO-US20731.  
XX PD

```

PR 28-JUL-1999; 99US-0146217.
XX
PA (GETH ) GENENTECH INC.
XX
XX Bodary SC, Fisher KL;
XX
XX WPI; 2001-182943/18.
XX
XX N-PSDB; AAF30377.
XX
XX Antibodies against ADAM8 polypeptides, useful e.g. for diagnosis and
XX treatment of tumours and inflammation.
XX
XX Disclosure; Page 5; 117pp; English.
XX
XX The present sequence is that of human ADAM8, a protein containing a
XX metalloprotease and a disintegrin domain. The ADAM8 gene (see
XX AAF30377) is amplified in the genome of cells of certain breast,
XX lung and colon cancers. Such gene amplification is associated with
XX the overexpression of the gene product and contributes to
XX tumorigenesis. Thus, the protein encoded by the amplified gene is
XX a useful target for the diagnosis and/or treatment or prevention of
XX certain cancers, and acts as a predictor of the prognosis of tumour
XX treatment. ADAM8 is believed to play a role in the migration of
XX monocytes through the extracellular matrix or possibly in the
XX processing of cytokines or other chemotactic molecules. Increased
XX ADAM8 expression in or near tumour cells may provide a means by
XX which tumour cells invade tissue or form metastases. Targeting such
XX a molecule with an antibody or another molecule that blocks,
XX inactivates or otherwise incapacitates the ability of tumour cells
XX to move into or out of tissue results in a clinically relevant
XX antitumour therapeutic. ADAM8 polypeptides can be used to raise
XX ADAM8-specific antibodies, and to screen for inhibitor compounds,
XX used to inhibit the growth of tumour cells.
XX
SQ Sequence 824 AA;
Query Match 47.9%; Score 46; DB 22; Length 824;
Best Local Similarity 43.8%; Pred. No. 1.1e+02;
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;
QY 2 KVIAPRRRLPDPQILKS 17
DB 30 EVVLPRRLPGPRVRA 45
RESULT 8
AAU86156.
ID AAU86156 standard; Protein; 824 AA.
XX
XX AAU86156;
XX
XX 15-JUL-2002 (first entry)
XX
XX Human.P01686 polypeptide.
XX
XX Human; PRO; benign tumour; malignant tumour; lymphoid malignancy;
XX leukaemia; neuronal disorder; stromal disorder; blastocoele disorder;
XX inflammatory disorder; immune disorder; angiogenic disorder;
XX cytostatic; neuroprotective.
XX
XX Homo sapiens.
XX
XX WO200153486-A1.
XX
XX 26-JUL-2001.
XX
XX 11-FEB-2000; 2000WO-US03565.
XX
XX 08-MAR-1999; 99WO-US05028.
XX
XX 11-MAR-1999; 99US-123972P.
XX
XX 11-MAY-1999; 99US-133459P.
XX
XX 02-JUN-1999; 99WO-US12252.
XX
XX 22-JUN-1999; 99US-140650P.
XX
PR 22-JUN-1999; 99US-140653P.
XX
XX 20-JUL-1999; 99US-144758P.
XX
XX 26-JUL-1999; 99US-145698P.
XX
XX 28-JUL-1999; 99US-146222P.
XX
XX 17-AUG-1999; 99US-149395P.
XX
XX 31-AUG-1999; 99US-151689P.
XX
XX 01-SEP-1999; 99WO-US20111.
XX
XX 15-SEP-1999; 99WO-US21090.
XX
XX 30-NOV-1999; 99WO-US28313.
XX
XX 01-DEC-1999; 99WO-US28301.
XX
XX 01-DEC-1999; 99WO-US28634.
XX
XX 05-JAN-2000; 2000WO-US00219.
XX
XX (GETH ) GENENTECH INC.
XX
XX Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Hillan KJ;
XX Marsters SA, Pan J, Pitti RM, Roy MA, Smith V, Stone DM;
XX Watanabe CK, Wood WI;
XX
XX WPI; 2002-205567/26.
XX
XX N-PSDB; ABK40282.
XX
XX Thirty five nucleic acids encoding PRO polypeptides, useful for
XX treating benign or malignant tumours, leukaemias and lymphoid
XX malignancies, inflammatory, angiogenic and immunologic disorders.
XX
XX Claim 61; Fig 58; 302pp; English.
XX
XX The present invention relates to the isolation of novel human PRO
XX polypeptides and the polynucleotide sequences encoding them. The
XX PRO polypeptides, agonists, antagonists or anti-PRO antibodies are
XX useful for treating benign or malignant tumours (e.g. renal, kidney,
XX bladder, breast, etc), leukaemias and lymphoid malignancies, other
XX disorders such as neuronal, glial, astrocytic, hypothalamic, glandular,
XX macrophagal, stromal and blastocoele disorders, inflammatory, immune
XX and angiogenic disorders. The polynucleotide sequences are also
XX useful in gene therapy. AAU86128-AAU86162 represent the human PRO
XX polypeptides of the invention.
XX
SQ Sequence 824 AA;
Query Match 47.9%; Score 46; DB 23; Length 824;
Best Local Similarity 43.8%; Pred. No. 1.1e+02;
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;
QY 2 KVIAPRRRLPDPQILKS 17
DB 30 EVVLPRRLPGPRVRA 45
RESULT 9
AAW85480
ID AAW85480 standard; Peptide; 19 AA.
XX
XX AAW85480;
XX
XX 15-MAR-1999 (first entry)
XX
XX Human GP88 autocrine growth factor antigenic peptide E19V.
XX
XX GP88; granulatin; epithelin; human; growth factor; autocrine; tumour;
XX cancer; viral infection; antagonist; therapy; diagnosis; antigen;
XX antibody.
XX
XX Homo sapiens.
XX
XX WO9852607-A1.
XX
XX 26-NOV-1998.
XX
XX 22-MAY-1998; 98WO-US10555.
XX
XX 16-DEC-1997; 97US-0991862.
XX

```

PR 23-MAY-1997; 97US-0863079.  
 PA (SERR/) SERRERO G.  
 XX Serrero G;  
 PI  
 DR WPI; 1999-045276/04.  
 XX Composition containing antagonist of growth factor GP88 - useful for  
 PT treating cancer and viral diseases and also for diagnosing disease  
 PT from altered GP88 expression  
 XX  
 PS Example 8; Page 45; 86pp; English.  
 XX This is the amino acid sequence of peptide E19V, comprising amino  
 CC acid residues E340-V364 of human GP88 (see AAW85475). GP88 is an 88  
 CC kDa glycoprotein autocrine growth factor that is expressed in a  
 CC tightly regulated manner in normal cells, is overexpressed and  
 CC unregulated in highly tumorigenic cells derived from normal cells,  
 CC and which acts as a stringently required growth stimulator for the  
 CC tumorigenic cells. E19V was used to raise neutralising antibodies  
 CC to GP88. Antagonists to GP88, such as anti-GP88 antibodies, are  
 CC used to treat diseases associated with increased expression of  
 CC GP88, particularly cancer but also viral infections. Anti-GP88  
 CC antibodies can also be used as diagnostic reagents and to deliver  
 CC toxins or other compounds to GP88-expressing cells.  
 XX  
 SQ Sequence 19 AA;  
 Query Match 45.8%; Score 44; DB 20; Length 19;  
 Best Local Similarity 55.6%; Pred. No. 4.4;  
 Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;  
 QY I KKVIAPRLPPQILKSD 18  
 : I I I I I I I I I  
 Db 1 EKAPAHLSLPDQALKRD 18  
 RESULT 10  
 AAW18298  
 ID AAW18298 standard; Peptide; 29 AA.  
 XX  
 AC AAW18298;  
 XX  
 DT 03-FEB-1998 (first entry)  
 XX  
 DE Human calpastatin B-cell epitope.  
 XX  
 KW Calpastatin; testis; human; calpain; protease inhibitor; immunogen;  
 KW vaccine; contraceptive; infertility; diagnosis; epitope;  
 KW B-cell; B-lymphocyte.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 4..17  
 XX /note= "Claim 18"  
 XX  
 PN WO9726001-Al.  
 XX  
 PD 24-JUL-1997.  
 XX  
 PF 15-JAN-1997; 97WO-US00908.  
 XX  
 PR 16-JAN-1996; 96US-0586592.  
 XX  
 PA (NOUN ) UNIV NORTHWESTERN.  
 XX  
 PI Goldberg E, Weinberg PO;  
 XX  
 DR WPI; 1997-385110/35.  
 XX  
 PT New testis-specific isoform of calpastatin - and derived immunogens,

PT useful in contraceptive vaccines for males or females, and for  
 PT diagnosing infertility  
 XX  
 PS Claim 17; Page 68; 86pp; English.  
 XX  
 CC This peptide comprises a B-cell epitope from a calpastatin isoform  
 CC (see AAW18394) found in human testis and sperm, but not in other  
 CC tissues. Calpastatin is an inhibitor of the cysteine protease  
 CC calpain. The B-cell epitope, or a shortened B-cell epitope (see  
 CC AAW19298), can be linked to a carrier comprising a universal T-cell  
 CC epitope from tetanus toxoid (see AAW18295 and AAW18296) to produce  
 CC an immunogen (see AAW18300). The B-cell epitope or immunogen can be  
 CC produced by chemical synthesis or as recombinant peptides in host  
 CC cells. They are useful as contraceptive vaccines that inhibit  
 CC fertilisation of an egg by sperm. The B-cell epitope can also  
 CC be used to detect specific antibodies in body fluids, the presence  
 CC of these antibodies being indicative of infertility.  
 XX  
 SQ Sequence 29 AA;  
 Query Match 45.8%; Score 44; DB 18; Length 29;  
 Best Local Similarity 44.4%; Pred. No. 6.8;  
 Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 KKVIAPRLPPQILKSD 18  
 : I I I I I I I I I  
 Db 6 EKKHPRRRPEPKIIPSE 23  
 RESULT 11  
 AAB94550  
 ID AAB94550 standard; Protein; 413 AA.  
 XX  
 AC AAB94550;  
 XX  
 DT 26-JUN-2001 (first entry)  
 XX  
 DE Human protein sequence SEQ ID NO:15310.  
 XX  
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP1074617-A2.  
 XX  
 PD 07-FEB-2001.  
 XX  
 PF 28-JUL-2000; 2000EP-0116126.  
 XX  
 PR 29-JUL-1999; 99JP-0248036.  
 PR 27-AUG-1999; 99JP-0300253.  
 PR 11-JAN-2000; 2000JP-0118776.  
 PR 02-MAY-2000; 2000JP-0183767.  
 PR 09-JUN-2000; 2000JP-0241899.  
 XX  
 PA (HELI-) HELIX RES INST.  
 XX  
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX  
 DR WPI; 2001-318749/34.  
 XX  
 PT Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -  
 XX  
 PS Claim 8; SEQ ID 15310; 2537pp + CD ROM; English.  
 XX  
 CC The present invention describes primer sets for synthesising 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of

CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to  
 CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.

XX SQ Sequence 413 AA;

Query Match 45.8%; Score 44; DB 22; Length 413;

Best Local Similarity 55.6%; Pred. No. 1.1e+02;

Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 1 KKVIAAPRLPDQILKSD 18  
 :| | | | | | | | | |  
 Db 166 EKAPAHLSLPDQALKRD 183

RESULT 12

AAW19394

ID AAW19394 standard; Protein; 463 AA.

XX AC AAW19394;

XX DT 03-FEB-1998 (first entry)

XX DE Human calpastatin (clone C-2 product);

XX KW Calpastatin; testis; human; calpain; protease inhibitor; immunogen;  
 KW vaccine; contraceptive; infertility; diagnosis; epitope.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Peptide 426...454

FT Peptide /note="B-cell epitope (Claim 17)"

FT Peptide 429...443

FT /note="B cell epitope (Claim 18)"

XX PN W09726001-A1.

XX PD 24-JUL-1997.

XX PF 16-JAN-1997; 97WO-US00908.

XX PR 16-JAN-1996; 96US-0586592.

XX PA (NOON ) UNIV NORTHWESTERN.

XX PI Goldberg E, Weinberg PO;

XX DR WPI; 1997-385110/35.

XX DR N-PSDB; AAT72643.

XX PT New testis-specific isoform of calpastatin - and derived immunogens,  
 PT useful in contraceptive vaccines for males or females, and for  
 PT diagnosing infertility

XX PS Example 16; Page 39-41; 86pp; English.

XX CC This polypeptide comprises an isoform of human calpastatin found

CC in sperm and testis but not in other tissues. It can be produced  
 CC in transformed host cells using isolated cDNA clone C-2 (see  
 CC AAT72643). A B-cell epitope (see AAW18298-99) of this calpastatin  
 CC isoform can be used to prepare an immunogen (see AAW18300) useful  
 CC as a male or female contraceptive vaccine that inhibits egg  
 CC fertilisation. The protein and B-cell epitope can also be used in  
 CC methods for detecting specific antibodies in body fluids, the  
 CC presence of these antibodies being indicative of infertility.

XX SQ Sequence 463 AA;

Query Match 45.8%; Score 44; DB 18; Length 463;

Best Local Similarity 44.4%; Pred. No. 1.2e+02;

Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 KKVIAAPRLPDQILKSD 18  
 :| | | | | | | | | |  
 Db 431 EKKHTPRRPEKIPSE 448

RESULT 13

AAW14326

ID AAW14326 standard; Protein; 593 AA.

XX AC AAW14326;

XX DT 17-JAN-1992 (first entry)

XX DE Human epithelin precursor.

XX KW ET; growth regulation; inhibition; stimulation.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Protein 1..593

FT /label= precursor

FT /note="claim 1, page 53"

FT Protein 282..337

FT /label= EP-1

FT /note="claim 2, page 53"

FT Protein 206..262

FT /label= EP-2

FT /note="claim 3, page 53"

FT Peptide 59..114

FT /label= EP

FT /note="claim 4, page 53"

FT Peptide 124..180

FT /label= EP

FT /note="claim 5, page 53"

FT Peptide 364..418

FT /label= EP

FT /note="claim 6, page 53"

FT Peptide 442..497

FT /label= EP

FT /note="claim 7, page 53"

FT Peptide 519..574

FT /label= EP

FT /note="claim 8, page 55"

XX PN W09115510-A.

XX PD 17-OCT-1991.

XX PF 03-APR-1991; 91WO-US02321.

XX PR 13-MAR-1991; 91US-0083796.

XX PR 03-APR-1990; 90US-0504508.

XX PA (BRIM.) BRISTOL-MYERS SQUIB.

XX PI Shoyab M, Plowman GD;

DR WPI: 1991-325168/44.  
 DR N-PSDB; AAQ14339.  
 XX  
 PT New cysteine-rich growth modulating proteins, epithelins - useful  
 PT as inhibitors of neoplastic cell growth and to promote wound  
 PT healing and treat psoriasis.  
 XX  
 PS Disclosure; Fig 22; 97pp; English.  
 XX  
 CC ET-1 is a bifunctional growth regulator, capable of stimulating  
 CC the growth of some cell types while inhibiting the growth of others.  
 CC ET-2 is functionally similar to ET-1 w.r.t. growth inhibitory  
 CC bioactivity. In contrast, however, ET-2 is apparently not capable of  
 CC eliciting the growth stimulatory activity characteristic of ET-1 and,  
 CC in fact, antagonises this ET-1 activity.  
 CC See also AAQ14338-40, AAQ14952-53, AAR14328-9 and AAR15315-20.  
 XX  
 SQ Sequence 593 AA;

Query Match 45.8%; Score 44; DB 12; Length 593;  
 Best Local Similarity 55.6%; Pred No. 1.6e+02;  
 Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 KKVIAPRRLPDQILKSD 18  
 :| | | | | | | |  
 Db 346 EKAPAHLSLPDQALKRD 363

RESULT 14  
 AAR48673  
 ID AAR48673 standard; Protein; 593 AA.  
 AC AAR48673;  
 XX  
 DT 22-APR-1994 (first entry)  
 XX  
 DE Granulin sequence.  
 XX  
 KW Granulin; keratinocytes; wound healing; inhibition; peptide;  
 KW granulocytes; leucocytes.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 340...364  
 FT /note= "E19V peptide used to raise antibody"  
 FT Peptide 566...579  
 FT /note= "A14R peptide used to raise antibody"  
 XX  
 PN WO9852607-A1.  
 XX  
 PD 26-NOV-1998.  
 XX  
 PF 22-MAY-1998; 98WO-US10555.  
 XX  
 PR 16-DEC-1997; 97US-0991862.  
 PR 23-MAY-1997; 97US-0863079.  
 XX  
 PA (SERR/) SERRERO G.  
 XX  
 PI Serrero G;  
 XX  
 DR WPI: 1999-045276/04.  
 DR N-PSDB; AAV82825.  
 XX  
 PT Composition containing antagonist of growth factor GP88 - useful for  
 PT treating cancer and viral diseases and also for diagnosing disease  
 PT from altered GP88 expression  
 XX  
 PS Example 5; Fig 9A; 86pp; English.  
 XX  
 CC This is the amino acid sequence of human GP88, an 88 kDa  
 CC glycoprotein autocrine growth factor and epithelin/granulin  
 CC precursor that is expressed in a tightly regulated manner in normal  
 CC cells, is overexpressed and unregulated in highly tumorigenic cells  
 CC derived from normal cells, and which acts as a stringently required  
 CC growth stimulator for the tumorigenic cells. Inhibition of GP88  
 CC expression or action in the tumorigenic cells results in an  
 CC inhibition of the tumorigenic properties of the overproducing  
 CC cells. Antagonists to GP88 are used to treat diseases associated  
 CC with increased expression of GP88, particularly cancer but also  
 CC viral infections. Fragments of GP88 are used to raise specific  
 CC antibodies (used as antagonists, as diagnostic reagents and for  
 CC delivering toxins or other compounds to GP88-expressing cells) and  
 CC to screen for antibodies. Methods are provided for diagnosing  
 CC disease, or determining susceptibility to disease, resulting from  
 CC altered GP88 activity.  
 XX  
 SQ Sequence 593 AA;

Query Match 45.8%; Score 44; DB 14; Length 593;  
 Best Local Similarity 55.6%; Pred No. 1.6e+02;  
 Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 KKVIAPRRLPDQILKSD 18  
 :| | | | | | | |  
 Db 346 EKAPAHLSLPDQALKRD 363

RESULT 15  
 AAW85475  
 ID AAW85475 standard; Protein; 593 AA.  
 XX  
 AC AAW85475;  
 XX  
 DT 15-MAR-1999 (first entry)  
 XX  
 DE Human GP88 autocrine growth factor.  
 XX  
 KW GP88; granulin; epithelin; human; growth factor; autocrine; tumour;  
 KW cancer; viral infection; antagonist; therapy; diagnosis.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 340...364  
 FT /note= "E19V peptide used to raise antibody"  
 FT Peptide 566...579  
 FT /note= "A14R peptide used to raise antibody"  
 XX  
 PN WO9852607-A1.  
 XX  
 PD 26-NOV-1998.  
 XX  
 PF 22-MAY-1998; 98WO-US10555.  
 XX  
 PR 16-DEC-1997; 97US-0991862.  
 PR 23-MAY-1997; 97US-0863079.  
 XX  
 PA (SERR/) SERRERO G.  
 XX  
 PI Serrero G;  
 XX  
 DR WPI: 1999-045276/04.  
 DR N-PSDB; AAV82825.  
 XX  
 PT Composition containing antagonist of growth factor GP88 - useful for  
 PT treating cancer and viral diseases and also for diagnosing disease  
 PT from altered GP88 expression  
 XX  
 PS Example 5; Fig 9A; 86pp; English.  
 XX  
 CC This is the amino acid sequence of human GP88, an 88 kDa  
 CC glycoprotein autocrine growth factor and epithelin/granulin  
 CC precursor that is expressed in a tightly regulated manner in normal  
 CC cells, is overexpressed and unregulated in highly tumorigenic cells  
 CC derived from normal cells, and which acts as a stringently required  
 CC growth stimulator for the tumorigenic cells. Inhibition of GP88  
 CC expression or action in the tumorigenic cells results in an  
 CC inhibition of the tumorigenic properties of the overproducing  
 CC cells. Antagonists to GP88 are used to treat diseases associated  
 CC with increased expression of GP88, particularly cancer but also  
 CC viral infections. Fragments of GP88 are used to raise specific  
 CC antibodies (used as antagonists, as diagnostic reagents and for  
 CC delivering toxins or other compounds to GP88-expressing cells) and  
 CC to screen for antibodies. Methods are provided for diagnosing  
 CC disease, or determining susceptibility to disease, resulting from  
 CC altered GP88 activity.  
 XX  
 SQ Sequence 593 AA;

Best Local Similarity 55.6%; Pred. No. 1.6e+02;  
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 KVIAPRRRLPDPQILKSD 18  
: | | | | | | | | | |  
Db 346 EKAPAHLSLDPDPQALKED 363

Search completed: July 8, 2003, 16:24:13  
Job time : 49.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2003, 16:21:00 ; Search time 11.6923 Seconds  
(without alignments)  
47.812 Million cell updates/sec

Title: US-09-824-647-3

Perfect score: 96

Sequence: 1 KKVAPRRRLDPQILKSDT 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/PTUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96	100.0	19	4	US-08-991-862-3
2	96	100.0	589	1	US-07-668-648-6
3	96	100.0	589	2	US-08-429-998-6
4	96	100.0	589	2	US-08-431-333-6
5	96	100.0	589	5	PCT-US91-02321-6
6	89	92.7	589	4	US-08-991-862-2
7	62	64.6	589	1	US-07-668-648-2
8	62	64.6	589	2	US-08-429-998-2
9	62	64.6	589	2	US-08-431-333-2
10	62	64.6	589	5	PCT-US91-02321-2
11	44	45.8	19	4	US-08-991-862-6
12	44	45.8	593	1	US-07-668-648-4
13	44	45.8	593	2	US-08-429-998-4
14	44	45.8	593	2	US-08-431-333-4
15	44	45.8	593	4	US-08-991-862-17
16	44	45.8	593	5	PCT-US91-02321-4
17	42	43.8	556	2	US-08-505-377-1
18	42	43.8	556	3	US-08-798-269-1
19	42	43.8	556	4	US-09-055-210-1
20	42	43.8	556	4	US-09-298-924-8
21	42	43.8	820	2	US-08-380-162-23
22	41	42.7	803	2	US-08-907-166-4
23	40	41.7	335	4	US-09-152-060-64
24	39	40.6	690	4	US-09-291-170A-2
25	39	40.6	690	4	US-09-724-884-2
26	38	39.6	97	4	US-09-152-060-99
27	38	39.6	99	4	US-09-336-536-34

28	38	39.6	143	4	US-09-152-060-84	Sequence 84, Appl
29	38	39.6	227	1	US-08-597-236-6	Sequence 6, Appl
30	38	39.6	227	1	US-08-746-682A-6	Sequence 6, Appl
31	38	39.6	280	2	US-08-464-517-6	Sequence 6, Appl
32	38	39.6	280	3	US-08-463-772-6	Sequence 6, Appl
33	38	39.6	291	5	PCT-US93-05000-6	Sequence 6, Appl
34	38	39.6	292	2	US-08-464-517-23	Sequence 23, Appl
35	38	39.6	292	2	US-08-246-361A-23	Sequence 23, Appl
36	38	39.6	292	3	US-08-463-772-23	Sequence 23, Appl
37	38	39.6	292	3	PCT-US93-05000-23	Sequence 23, Appl
38	38	39.6	292	5	US-09-336-536-29	Sequence 29, Appl
39	38	39.6	341	4	US-09-336-536-28	Sequence 28, Appl
40	38	39.6	370	4	US-08-979-424-1	Sequence 1, Appl
41	38	39.6	390	2	US-08-706-216-6	Sequence 6, Appl
42	38	39.6	391	4	US-08-790-186A-4	Sequence 4, Appl
43	38	39.6	422	4	US-08-975-763-69	Sequence 69, Appl
44	38	39.6	464	4	US-09-295-028-69	Sequence 69, Appl
45	38	39.6	464	4		

ALIGNMENTS

RESULT 1  
US-08-991-862-3  
; Sequence 3, Application US/08991862.  
; Patent No. 6309826  
; GENERAL INFORMATION:  
; APPLICANT: Serrero, Ginette  
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
; FILE REFERENCE: 29996.488/P001-A  
; CURRENT APPLICATION NUMBER: US/08/991,862  
; CURRENT FILING DATE: 1998-08-17  
; EARLIER APPLICATION NUMBER: 08/863,862  
; EARLIER FILING DATE: 1997-05-23  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: Patent Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: mouse granulin  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(19)  
; OTHER INFORMATION: Internal peptide of mouse GP88 used to raise the  
; OTHER INFORMATION: antiserum against the GP88 used in the  
; OTHER INFORMATION: immunoaffinity step.  
US-08-991-862-3

Query Match 100.0%; Score 96; DB 4; Length 19;  
Best Local Similarity 100.0%; Pred. No. 1.3e-09;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKVAPRRRLDPQILKSDT 19  
DB 1 KKVAPRRRLDPQILKSDT 19

RESULT 2  
US-07-668-648-6  
; Sequence 6, Application US/07668648  
; Patent No. 5416192  
; GENERAL INFORMATION:  
; APPLICANT: Shoyab, Mohammed  
; APPLICANT: Plowman, Gregory D.  
; TITLE OF INVENTION: EPITHELIAL NOVEL CYSTEINE-RICH GROWTH  
; NUMBER OF SEQUENCES: 12  
; TITLE OF INVENTION: MODULATING PROTEINS  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York

COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/668,648  
FILING DATE: 19910819  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Misrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELEPHONE: (212)790-9090  
TELEFAX: (212) 869-9741  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 589 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-668-648-6

Query Match 100.0%; Score 96; DB 1; Length 589;  
Best Local Similarity 100.0%; Pred. No. 6.3e-08;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKVAPRRRLPDPQILKSDT 19  
DB 344 KKVAPRRRLPDPQILKSDT 362

RESULT 3  
US-08-429-998-6  
Sequence 6, Application US/08429998  
Patent No. 5885961  
GENERAL INFORMATION:  
APPLICANT: Shoyab, Mohammed  
APPLICANT: Plowman, Gregory D.  
TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH  
TITLE OF INVENTION: MODULATING PROTEINS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/429,998  
FILING DATE: 27-APR-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/668,648  
FILING DATE: 13-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Misrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELEPHONE: (212)790-9090  
TELEFAX: (212) 869-9741  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:

LENGTH: 589 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-429-998-6

Query Match 100.0%; Score 96; DB 2; Length 589;  
Best Local Similarity 100.0%; Pred. No. 6.3e-08;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKVAPRRRLPDPQILKSDT 19  
DB 344 KKVAPRRRLPDPQILKSDT 362

RESULT 4  
US-08-431-333-6  
Sequence 6, Application US/08431333  
Patent No. 5965723  
GENERAL INFORMATION:  
APPLICANT: Shoyab, Mohammed  
APPLICANT: Plowman, Gregory D.  
TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH  
TITLE OF INVENTION: MODULATING PROTEINS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/431,333  
FILING DATE: 27-APR-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/668,648  
FILING DATE: 13-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Misrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELEPHONE: (212)790-9090  
TELEFAX: (212) 869-9741  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 589 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-431-333-6

Query Match 100.0%; Score 96; DB 2; Length 589;  
Best Local Similarity 100.0%; Pred. No. 6.3e-08;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKVAPRRRLPDPQILKSDT 19  
DB 344 KKVAPRRRLPDPQILKSDT 362

RESULT 5  
PCT-US91-02321-6  
Sequence 6, Application PC/TUS9102321  
GENERAL INFORMATION:  
APPLICANT: Shoyab, Mohammed



APPLICANT: Plowman, Gregory D.  
TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH  
MODULATING PROTEINS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Bristol-Myers Squibb Company  
STREET: 3005 First Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98121  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/02321  
FILING DATE: 19910403  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Poor, Brian W.  
REGISTRATION NUMBER: 32,928  
REFERENCE/DOCKET NUMBER: ON0071A-PC  
TELEPHONE: (206)728-4800  
TELEFAX: (206)448-4775  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 589 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US91-02321-6

Query Match 100.0%; Score 96; DB 5; Length 589;  
Best Local Similarity 100.0%; Pred. No. 6.3e-08;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKVAPRRLPDPQILKSDT 19  
DB 344 KKVAPRRLPDPQILKSDT 362

## RESULT 6

US-08-991-862-2  
Sequence 2, Application US/08991862  
Patent No. 6309826  
GENERAL INFORMATION:  
APPLICANT: Serrero, GINETTE  
FILE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
CURRENT APPLICATION NUMBER: US/08/991,862  
CURRENT FILING DATE: 1998-08-17  
EARLIER APPLICATION NUMBER: 08/863,862  
EARLIER FILING DATE: 1997-05-23  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 2  
LENGTH: 589  
TYPE: PRT  
ORGANISM: Mouse epithelin/granulin  
US-08-991-862-2

Query Match 92.7%; Score 89; DB 4; Length 589;  
Best Local Similarity 94.7%; Pred. No. 9e-07;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKVAPRRLPDPQILKSDT 19  
DB 344 KKVAPRRLPDPQILKSDT 362

RESULT 7  
US-07-668-648-2  
Sequence 2, Application US/07668648  
Patent No. 5416192  
GENERAL INFORMATION:  
APPLICANT: Shoyab, Mohammed  
APPLICANT: Plowman, Gregory D.  
TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH  
MODULATING PROTEINS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/668,648  
FILING DATE: 19910819  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Misrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELEPHONE: (212)790-9090  
TELEFAX: (212) 869-9741  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 589 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-668-648-2

Query Match 64.6%; Score 62; DB 1; Length 589;  
Best Local Similarity 72.2%; Pred. No. 0.025;  
Matches 13; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KKVAPRRLPDPQILKSD 18  
DB 344 KKVAPRRLPDPQILKSD 361

## RESULT 8

US-08-429-998-2  
Sequence 2, Application US/08429998  
Patent No. 5885961  
GENERAL INFORMATION:  
APPLICANT: Shoyab, Mohammed  
APPLICANT: Plowman, Gregory D.  
TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH  
MODULATING PROTEINS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/429,998  
FILING DATE: 27-APR-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/668,648  
FILING DATE: 13-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Misrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELEPHONE: (212)790-9090  
TELEFAX: (212) 869-9741  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 589 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-429-998-2

Query Match 64.6%; Score 62; DB 2; Length 589;  
Best Local Similarity 72.2%; Pred. No. 0.025;  
Matches 13; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KKVAPRRRLPDPQILKSD 18  
DB 344 KKVATASLPLDPQILKND 361

RESULT 9  
US-08-431-333-2  
Sequence 2, Application US/08431333  
Patent No. 5965723  
GENERAL INFORMATION:  
APPLICANT: Shoyab, Mohammed  
TITLE OF INVENTION: EPIITHELINS: NOVEL CYSTEINE-RICH GROWTH  
MODULATING PROTEINS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/431,333  
FILING DATE: 27-APR-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/668,648  
FILING DATE: 13-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Misrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELEPHONE: (212)790-9090  
TELEFAX: (212) 869-9741  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 589 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-431-333-2

Query Match 64.6%; Score 62; DB 2; Length 589;  
Best Local Similarity 72.2%; Pred. No. 0.025;  
Matches 13; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KKVAPRRRLPDPQILKSD 18  
DB 344 KKVATASLPLDPQILKND 361

RESULT 10  
PCT-US91-02321-2  
Sequence 2, Application PC/TUS9102321  
GENERAL INFORMATION:  
APPLICANT: Shoyab, Mohammed  
TITLE OF INVENTION: EPIITHELINS: NOVEL CYSTEINE-RICH GROWTH  
MODULATING PROTEINS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Bristol-Myers Squibb Company  
STREET: 3005 First Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98121  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/02321  
FILING DATE: 19910403  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Poor, Brian W.  
REGISTRATION NUMBER: 32,928  
REFERENCE/DOCKET NUMBER: ON0071A-PC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)728-4800  
TELEFAX: (206)448-4775  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 589 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US91-02321-2

Query Match 64.6%; Score 62; DB 5; Length 589;  
Best Local Similarity 72.2%; Pred. No. 0.025;  
Matches 13; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KKVAPRRRLPDPQILKSD 18  
DB 344 KKVATASLPLDPQILKND 361

RESULT 11  
US-08-991-862-6  
Sequence 6, Application US/08991862  
Patent No. 6309826  
GENERAL INFORMATION:  
APPLICANT: Serrero, Ginette  
TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
FILE REFERENCE: Z9996.488/P001-A  
CURRENT APPLICATION NUMBER: US/08/991,862  
CURRENT FILING DATE: 1998-08-17  
EARLIER APPLICATION NUMBER: 08/863,862  
EARLIER FILING DATE: 1997-05-23  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: PatentIn Ver. 2.0

```
; SEQ ID NO 6
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Human granulin
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(19)
; OTHER INFORMATION: Internal peptide of human GP88 used to develop
; OTHER INFORMATION: neutralizing anti-human GP88 monoclonal antibody.
US-08-991-862-6

Query Match 45.8%; Score 44; DB 4; Length 19;
Best Local Similarity 55.6%; Pred. No. 0.45;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 KKVAPRRLLPDPQILKSD 18
: | | | | | | | | |
Db 1 EKAPAHLSLPDPQALKRD 18

RESULT 12
US-07-668-648-4
; Sequence 4, Application US/07668648
; Patent No. 5416192
; GENERAL INFORMATION:
; APPLICANT: Shoyab, Mohammed
; APPLICANT: Plovman, Gregory D.
; TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH
; TITLE OF INVENTION: MODULATING PROTEINS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/668.648
; FILING DATE: 19910819
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 5624-161-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)790-9090
; TELEFAX: (212) 869-9741
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 593 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-668-648-4

Query Match 45.8%; Score 44; DB 1; Length 593;
Best Local Similarity 55.6%; Pred. No. 23;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 KKVAPRRLLPDPQILKSD 18
: | | | | | | | | |
Db 346 EKAPAHLSLPDPQALKRD 363

RESULT 13
US-08-429-998-4
; Sequence 4, Application US/08429998
; Patent No. 5416192
; GENERAL INFORMATION:
; APPLICANT: Shoyab, Mohammed
; APPLICANT: Plovman, Gregory D.
; TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH
; TITLE OF INVENTION: MODULATING PROTEINS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/429,998
; FILING DATE: 27-APR-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/668,648
; FILING DATE: 13-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 5624-161-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)790-9090
; TELEFAX: (212) 869-9741
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 593 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-429-998-4

Query Match 45.8%; Score 44; DB 2; Length 593;
Best Local Similarity 55.6%; Pred. No. 23;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 KKVAPRRLLPDPQILKSD 18
: | | | | | | | | |
Db 346 EKAPAHLSLPDPQALKRD 363

RESULT 14
US-08-431-333-4
; Sequence 4, Application US/08431333
; Patent No. 5965723
; GENERAL INFORMATION:
; APPLICANT: Shoyab, Mohammed
; APPLICANT: Plovman, Gregory D.
; TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH
; TITLE OF INVENTION: MODULATING PROTEINS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```

APPLICATION NUMBER: US/08/431,333  
FILING DATE: 27-APR-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/668,648  
FILING DATE: 13-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)790-9090  
TELEFAX: (212) 869-9741  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 593 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-431-333-4

Query Match 45.8%; Score 44; DB 2; Length 593;  
Best Local Similarity 55.6%; Pred. No. 23;  
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 KKVIAPRRLPDPOILKSD 18  
DB 346 EKAPAHLSLPDQALKRD 363

RESULT 15  
US-08-991-862-17  
Sequence 17, Application US/08991862  
Patent No. 6309826  
GENERAL INFORMATION:  
APPLICANT: Seritiero, Ginette  
TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
FILE REFERENCE: 29996.488/P001-A  
CURRENT APPLICATION NUMBER: US/08/991,862  
CURRENT FILING DATE: 1998-08-17  
EARLIER APPLICATION NUMBER: 08/863,862  
EARLIER FILING DATE: 1997-05-23  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 17  
LENGTH: 593  
TYPE: PRT  
ORGANISM: Human GP88 cDNA  
US-08-991-862-17

Query Match 45.8%; Score 44; DB 4; Length 593;  
Best Local Similarity 55.6%; Pred. No. 23;  
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 KKVIAPRRLPDPOILKSD 18  
DB 346 EKAPAHLSLPDQALKRD 363

Search completed: July 8, 2003, 16:30:16  
Job time : 13.6923 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2003, 16:24:26 ; Search time 18.5128 Seconds  
(without alignments)  
119.483 Million cell updates/sec

Title: US-09-824-647-3  
Perfect score: 96  
Sequence: 1 KKVAPRLPPDQILKSDT 19

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 116419773 residues

Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA:  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96	100.0	19	US-09-824-647-3	Sequence 3, Appli
2	96	100.0	19	US-10-218-509-3	Sequence 3, Appli
3	96	100.0	19	US-10-281-160-3	Sequence 3, Appli
4	96	100.0	19	US-09-813-156-3	Sequence 3, Appli
5	96	100.0	19	US-09-824-807-3	Sequence 3, Appli
6	89	92.7	589	US-09-824-647-2	Sequence 2, Appli
7	89	92.7	589	US-10-218-509-2	Sequence 2, Appli
8	89	92.7	589	US-10-281-160-2	Sequence 2, Appli
9	89	92.7	589	US-09-813-156-2	Sequence 2, Appli
10	89	92.7	589	US-09-824-807-2	Sequence 2, Appli
11	47	49.0	92	US-09-800-971-11	Sequence 11, Appli
12	46	47.9	824	US-10-226-844-1	Sequence 1, Appli
13	44	45.8	19	US-09-824-647-6	Sequence 6, Appli
14	44	45.8	19	US-10-218-509-6	Sequence 6, Appli
15	44	45.8	19	US-10-281-160-6	Sequence 6, Appli
16	44	45.8	19	US-09-813-156-6	Sequence 6, Appli
17	44	45.8	19	US-09-824-807-6	Sequence 6, Appli
18	44	45.8	319	US-10-156-761-1112	Sequence 1112, A
19	44	45.8	593	US-09-824-647-17	Sequence 17, Appli

Sequence 17, Appli  
Sequence 17, Appli  
Sequence 17, Appli  
Sequence 17, Appli  
Sequence 1416, Ap  
Sequence 8, Appli  
Sequence 13, Appli  
Sequence 5214, Ap  
Sequence 6120, Ap  
Sequence 5837, Ap  
Sequence 46, Appli  
Sequence 48, Appli  
Sequence 178, Appli  
Sequence 2, Appli  
Sequence 4, Appli  
Sequence 4, Appli  
Sequence 14, Appli  
Sequence 16, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 12754, A  
Sequence 19, Appli  
Sequence 64, Appli  
Sequence 64, Appli  
Sequence 64, Appli

20 44 45.8 593 9 US-10-218-509-17  
21 44 45.8 593 9 US-10-281-160-17  
22 44 45.8 593 10 US-09-813-156-17  
23 44 45.8 593 10 US-09-824-807-17  
24 44 45.8 621 10 US-09-925-301-1416  
25 43 44.8 596 10 US-09-797-039-8  
26 42 43.8 416 10 US-09-430-221-13  
27 41.5 43.2 450 10 US-09-815-242-5214  
28 41 42.7 362 9 US-09-738-626-6120  
29 41 42.7 429 9 US-09-738-626-5837  
30 41 42.7 666 9 US-09-746-660A-46  
31 41 42.7 798 9 US-09-533-029-48  
32 41 42.7 798 9 US-09-934-455-178  
33 41 42.7 803 9 US-10-029-382-2  
34 41 42.7 803 10 US-09-391-340-4  
35 41 42.7 803 10 US-09-948-369-4  
36 41 42.7 803 10 US-09-948-369-14  
37 41 42.7 803 10 US-09-948-369-16  
38 41 42.7 803 12 US-10-034-849-2  
39 41 42.7 803 12 US-10-034-621-2  
40 40.5 42.2 240 9 US-10-156-761-12754  
41 40 41.7 271 10 US-09-883-060-5  
42 40 41.7 321 9 US-10-186-886-19  
43 40 41.7 335 9 US-09-852-797-64  
44 40 41.7 335 10 US-09-853-161-64  
45 40 41.7 335 10 US-09-852-659A-64

## ALIGNMENTS

RESULT 1  
US-09-824-647-3  
; Sequence 3, Application US/09824647  
; Publication No. US20020183270A1  
; GENERAL INFORMATION:  
; APPLICANT: Serrero, Ginette  
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
; FILE REFERENCE: Z9996.488/P001-A  
; CURRENT APPLICATION NUMBER: US/09/824,647  
; CURRENT FILING DATE: 2001-04-04  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/991,862  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/863,862  
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-05-23  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: mouse granuln  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)...(19)  
; OTHER INFORMATION: Internal peptide of mouse GP88 used to raise the  
; OTHER INFORMATION: antisera against the GP88 used in the  
; OTHER INFORMATION: immunoaffinity step.  
US-09-824-647-3

Query Match 100.0% Score 96; DB 9; Length 19;  
Best Local Similarity 100.0%; Pred. No. 8.9e-09;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KKVAPRLPPDQILKSDT 19  
|||||  
Db 1 KKVAPRLPPDQILKSDT 19

RESULT 2  
US-10-218-509-3  
; Sequence 3, Application US/10218509  
; Publication No. US20030092661A1  
; GENERAL INFORMATION:

APPLICANT: Serrero, Ginette  
TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
FILE REFERENCE: Z9996.488/P001-A  
CURRENT APPLICATION NUMBER: US/10/218,509  
CURRENT FILING DATE: 2002-08-15  
PRIOR APPLICATION NUMBER: 08/991,862  
PRIOR FILING DATE: 1998-08-17  
PRIOR APPLICATION NUMBER: 08/863,862  
PRIOR FILING DATE: 1997-05-23  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 3  
LENGTH: 19  
TYPE: PRT  
ORGANISM: mouse granulin  
FEATURE:  
NAME/KEY: PEPTIDE  
LOCATION: (1)..(19)  
OTHER INFORMATION: Internal peptide of mouse GP88 used to raise the  
OTHER INFORMATION: antisera against the GP88 used in the  
OTHER INFORMATION: immunoaffinity step.  
US-10-218-509-3

Query Match 100.0%; Score 96; DB 9; Length 19;  
Best Local Similarity 100.0%; Pred. No. 8.9e-09;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKVAPRRRLDPQILKSDT 19  
DB 1 KKVAPRRRLDPQILKSDT 19

## RESULT 3

US-10-281-160-3  
Sequence 3, Application US/10281160  
Publication No. US20030108950A1

## GENERAL INFORMATION:

APPLICANT: Serrero, Ginette  
TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
FILE REFERENCE: Z9996.488/P001-A  
CURRENT APPLICATION NUMBER: US/10/281,160  
CURRENT FILING DATE: 2002-10-28  
PRIOR APPLICATION NUMBER: 08/991,862  
PRIOR FILING DATE: 1998-08-17  
PRIOR APPLICATION NUMBER: 08/863,862  
PRIOR FILING DATE: 1997-05-23  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 3  
LENGTH: 19  
TYPE: PRT  
ORGANISM: mouse granulin  
FEATURE:  
NAME/KEY: PEPTIDE  
LOCATION: (1)..(19)  
OTHER INFORMATION: Internal peptide of mouse GP88 used to raise the  
OTHER INFORMATION: antisera against the GP88 used in the  
OTHER INFORMATION: immunoaffinity step.  
US-10-281-160-3

Query Match 100.0%; Score 96; DB 9; Length 19;  
Best Local Similarity 100.0%; Pred. No. 8.9e-09;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKVAPRRRLDPQILKSDT 19  
DB 1 KKVAPRRRLDPQILKSDT 19

## RESULT 4

US-09-813-156-3  
Sequence 3, Application US/09813156  
Patent No. US20020061859A1

GENERAL INFORMATION:  
APPLICANT: Serrero, Ginette  
TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
FILE REFERENCE: Z9996.488/P001-A  
CURRENT APPLICATION NUMBER: US/09/813,156  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 08/991,862  
PRIOR FILING DATE: 1997-12-16  
PRIOR APPLICATION NUMBER: 08/863,862  
PRIOR FILING DATE: 1997-05-23  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 3  
LENGTH: 19  
TYPE: PRT  
ORGANISM: mouse granulin  
FEATURE:  
NAME/KEY: PEPTIDE  
LOCATION: (1)..(19)  
OTHER INFORMATION: Internal peptide of mouse GP88 used to raise the  
OTHER INFORMATION: antisera against the GP88 used in the  
OTHER INFORMATION: immunoaffinity step.  
US-09-813-156-3

Query Match 100.0%; Score 96; DB 10; Length 19;  
Best Local Similarity 100.0%; Pred. No. 8.9e-09;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKVAPRRRLDPQILKSDT 19  
DB 1 KKVAPRRRLDPQILKSDT 19

## RESULT 5

US-09-824-807-3  
Sequence 3, Application US/09824807  
Patent No. US20020094966A1

## GENERAL INFORMATION:

APPLICANT: Serrero, Ginette  
TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
FILE REFERENCE: Z9996.488/P001-A  
CURRENT APPLICATION NUMBER: US/09/824,807  
CURRENT FILING DATE: 2001-04-04  
PRIOR APPLICATION NUMBER: 08/991,862  
PRIOR FILING DATE: 1997-12-16  
PRIOR APPLICATION NUMBER: 08/863,862  
PRIOR FILING DATE: 1997-05-23  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 3  
LENGTH: 19  
TYPE: PRT  
ORGANISM: mouse granulin  
FEATURE:  
NAME/KEY: PEPTIDE  
LOCATION: (1)..(19)  
OTHER INFORMATION: Internal peptide of mouse GP88 used to raise the  
OTHER INFORMATION: antisera against the GP88 used in the  
OTHER INFORMATION: immunoaffinity step.  
US-09-824-807-3

Query Match 100.0%; Score 96; DB 10; Length 19;  
Best Local Similarity 100.0%; Pred. No. 8.9e-09;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKVAPRRRLDPQILKSDT 19  
DB 1 KKVAPRRRLDPQILKSDT 19

## RESULT 6

US-09-824-647-2  
Sequence 2, Application US/09824647

```
; Publication No. US20020183270A1
; GENERAL INFORMATION:
; APPLICANT: Serrero, Ginette
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
; FILE REFERENCE: Z9996.488/P001-A
; CURRENT APPLICATION NUMBER: US/09/824,647
; CURRENT FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/991,862
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/863,862
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-05-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Mouse epithelin/granulin
US-09-824-647-2

Query Match          92.7%; Score 89; DB 9; Length 589;
Best Local Similarity 94.7%; Pred. No. 6.1e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 KKVIAPRLPDPQILKSDT 19
        ||||| ||||| ||||| |||||
DB      344 KKVIAPRLPDPQILKSDT 362

RESULT 7
US-10-218-509-2
; Sequence 2, Application US/10218509
; Publication No. US20030092661A1
; GENERAL INFORMATION:
; APPLICANT: Serrero, Ginette
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
; FILE REFERENCE: Z9996.488/P001-A
; CURRENT APPLICATION NUMBER: US/10/218,509
; CURRENT FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: 08/991,862
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 08/863,862
; PRIOR FILING DATE: 1997-05-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Mouse epithelin/granulin
US-10-218-509-2

Query Match          92.7%; Score 89; DB 9; Length 589;
Best Local Similarity 94.7%; Pred. No. 6.1e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 KKVIAPRLPDPQILKSDT 19
        ||||| ||||| ||||| |||||
DB      344 KKVIAPRLPDPQILKSDT 362

RESULT 8
US-10-281-160-2
; Sequence 2, Application US/10281160
; Publication No. US20030108950A1
; GENERAL INFORMATION:
; APPLICANT: Serrero, Ginette
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
; FILE REFERENCE: Z9996.488/P001-A
; CURRENT APPLICATION NUMBER: US/10/281,160
; CURRENT FILING DATE: 2002-10-28
; PRIOR APPLICATION NUMBER: US/08/991,862
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 08/863,862
; PRIOR FILING DATE: 1997-05-23
```

```
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Mouse epithelin/granulin
US-10-281-160-2

Query Match          92.7%; Score 89; DB 9; Length 589;
Best Local Similarity 94.7%; Pred. No. 6.1e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 KKVIAPRLPDPQILKSDT 19
        ||||| ||||| ||||| |||||
DB      344 KKVIAPRLPDPQILKSDT 362

RESULT 9
US-09-813-156-2
; Sequence 2, Application US/09813156
; Patent No. US20020061859A1
; GENERAL INFORMATION:
; APPLICANT: Serrero, Ginette
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
; FILE REFERENCE: Z9996.488/P001-A
; CURRENT APPLICATION NUMBER: US/09/813,156
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 08/991,862
; PRIOR FILING DATE: 1997-12-16
; PRIOR APPLICATION NUMBER: 08/863,862
; PRIOR FILING DATE: 1997-05-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Mouse epithelin/granulin
US-09-813-156-2

Query Match          92.7%; Score 89; DB 10; Length 589;
Best Local Similarity 94.7%; Pred. No. 6.1e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 KKVIAPRLPDPQILKSDT 19
        ||||| ||||| ||||| |||||
DB      344 KKVIAPRLPDPQILKSDT 362

RESULT 10
US-09-824-807-2
; Sequence 2, Application US/09824807
; Patent No. US20020094966A1
; GENERAL INFORMATION:
; APPLICANT: Serrero, Ginette
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
; FILE REFERENCE: Z9996.488/P001-A
; CURRENT APPLICATION NUMBER: US/09/824,807
; CURRENT FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 08/991,862
; PRIOR FILING DATE: 1997-12-16
; PRIOR APPLICATION NUMBER: 08/863,862
; PRIOR FILING DATE: 1997-05-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Mouse epithelin/granulin
US-09-824-807-2

Query Match          92.7%; Score 89; DB 10; Length 589;
Best Local Similarity 94.7%; Pred. No. 6.1e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

QY 1 KVIAPRRRLPDQILKSDT 19  
||||| ||||| |||||  
Db 344 KVIAPRLRLPDQILKSDT 362

RESULT 11  
US-09-800-971-11  
; Sequence 11, Application US/09800971  
; Patent No. US20020098577A1  
; GENERAL INFORMATION:  
; APPLICANT: Rachel A. Meyers  
; TITLE OF INVENTION: 16835, A NOVEL HUMAN PHOSPHOLIPASE C  
; FILE REFERENCE: 10448-023001  
; CURRENT APPLICATION NUMBER: US/09/800,971  
; PRIOR FILING DATE: 2001-03-06  
; PRIOR APPLICATION NUMBER: 60/187,453  
; PRIOR FILING DATE: 2000-03-07  
; PRIOR APPLICATION NUMBER: 60/188,032  
; PRIOR FILING DATE: 2000-03-09  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 92  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: consensus sequence  
US-09-800-971-11

Query Match 49.0%; Score 47; DB 10; Length 92;  
Best Local Similarity 50.0%; Pred. No. 3;  
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 KVIAPRRRLPDQILKSDT 19  
||: ||||| ||: ||  
Db 5 KIISARNLPDPVKVSKT 22

RESULT 12  
US-10-226-844-1  
; Sequence 1, Application US/10226844  
; Publication No. US20030113764A1  
; GENERAL INFORMATION:  
; APPLICANT: Bodary, Sarah C.  
; APPLICANT: Fisher, Karen L.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF TUMORS  
; FILE REFERENCE: P1773R1  
; CURRENT APPLICATION NUMBER: US/10/226,844  
; PRIOR FILING DATE: 2002-08-22  
; PRIOR APPLICATION NUMBER: US/09/627,202  
; PRIOR FILING DATE: 2000-07-27  
; PRIOR APPLICATION NUMBER: US 60/146,217  
; PRIOR FILING DATE: 1999-07-28  
; NUMBER OF SEQ ID NOS: 22  
; SEQ ID NO 1  
; LENGTH: 824  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-10-226-844-1

Query Match 47.9%; Score 46; DB 9; Length 824;  
Best Local Similarity 43.8%; Pred. No. 55;  
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 2 KVIAPRRRLPDQILKS 17  
||: ||||| ||: ||  
Db 30 EVVLPRLPCPRVRA 45

RESULT 13  
US-09-824-647-6

; Sequence 6, Application US/09824647  
; Publication No. US20020183270A1  
; GENERAL INFORMATION:  
; APPLICANT: Serrero, Ginette  
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
; FILE REFERENCE: Z9996.488/P001-A  
; CURRENT APPLICATION NUMBER: US/09/824,647  
; CURRENT FILING DATE: 2001-04-04  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/991,862  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/863,862  
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-05-23  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Human granulin  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)-(19)  
; OTHER INFORMATION: Internal peptide of human GP88 used to develop  
; OTHER INFORMATION: neutralizing anti-human GP88 monoclonal antibody.  
US-09-824-647-6

Query Match 45.8%; Score 44; DB 9; Length 19;  
Best Local Similarity 55.6%; Pred. No. 1.4;  
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 KVIAPRRRLPDQILKSD 18  
:| | ||||| || |  
Db 1 EKAPAHLSLPDQALKRD 18

RESULT 14  
US-10-218-509-6  
; Sequence 6, Application US/10218509  
; Publication No. US20030092661A1  
; GENERAL INFORMATION:  
; APPLICANT: Serrero, Ginette  
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
; FILE REFERENCE: Z9996.488/P001-A  
; CURRENT APPLICATION NUMBER: US/10/218,509  
; CURRENT FILING DATE: 2002-08-15  
; PRIOR APPLICATION NUMBER: 08/991,862  
; PRIOR FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: 08/863,862  
; PRIOR FILING DATE: 1997-05-23  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Human granulin  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)-(19)  
; OTHER INFORMATION: Internal peptide of human GP88 used to develop  
; OTHER INFORMATION: neutralizing anti-human GP88 monoclonal antibody.  
US-10-218-509-6

Query Match 45.8%; Score 44; DB 9; Length 19;  
Best Local Similarity 55.6%; Pred. No. 1.4;  
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 KVIAPRRRLPDQILKSD 18  
:| | ||||| || |  
Db 1 EKAPAHLSLPDQALKRD 18

RESULT 15  
US-10-281-160-6  
; Sequence 6, Application US/10281160



; Publication No. US20030108950A1  
; GENERAL INFORMATION:  
; APPLICANT: Seriero, Ginette  
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
; FILE REFERENCE: 29996.488/P001-A  
; CURRENT APPLICATION NUMBER: US/10/281,160  
; CURRENT FILING DATE: 2002-10-28  
; PRIOR APPLICATION NUMBER: US/08/991,862  
; PRIOR FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: 08/863,862  
; PRIOR FILING DATE: 1997-05-23  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Human granulin  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(19)  
; OTHER INFORMATION: Internal peptide of human GP88 used to develop  
; OTHER INFORMATION: neutralizing anti-human GP88 monoclonal antibody.  
US-10-281-160-6

Query Match 45.8%; Score 44; DB 9; Length 19;  
Best Local Similarity 55.6%; Pred. No. 1.4;  
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 KVIAPRRLPDQILKSD 18  
Db 1 EKAPAHLSLPDQALKRD 18

Search completed: July 8, 2003, 16:31:39  
Job time : 19.5128 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2003, 16:19:45 ; Search time 19.2436 seconds  
(without alignments)  
94.918 Million cell updates/sec

Title: US-09-824-647-3

Perfect score: 96

Sequence: 1 KKVAPRRLPDPQILKSDT 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_73.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96	100.0	589	2 C38128	epithelin/granulin
2	62	64.6	589	2 B38128	epithelin/granulin
3	48	50.0	328	2 F81317	UDPglucose 4-epime
4	44.5	46.4	256	2 F83221	beta-ketoacyl redu
5	44	45.8	593	1 GYHU	granulin precursor
6	43	44.8	174	2 T07099	late embryogenesis
7	43	44.8	223	2 T00150	hypothetical prote
8	43	44.8	344	2 H87710	hypothetical prote
9	43	44.8	560	2 S56142	calcium-binding pr
10	43	44.8	753	2 T01619	hypothetical prote
11	42	43.8	244	2 S74911	hypothetical prote
12	42	43.8	280	2 D70987	probable plcb prot
13	42	43.8	556	2 JC5132	alpha-amylase (EC
14	42	43.8	566	2 D70709	hypothetical prote
15	42	43.8	820	1 D5ECK	thra bifunctional
16	42	43.8	820	2 B90629	aspartokinase I-ho
17	42	43.8	820	2 B85480	aspartokinase I, h
18	42	43.8	820	2 AC0502	aspartokinase I/ho
19	42	43.8	874	2 B70914	probable dnaB prot
20	42	43.8	1650	2 S28721	hypothetical prote
21	41.5	43.2	450	2 E82973	conserved hypoteth
22	41.5	43.2	622	2 D44986	apical membrane an
23	41.5	43.2	622	2 B44986	apical membrane an
24	41.5	43.2	622	2 A32499	apical membrane an
25	41	42.7	351	2 G82541	UDP-N-acetylpyruvo
26	41	42.7	368	2 T51200	hypothetical prote
27	41	42.7	378	1 D69583	alcohol dehydrogen
28	41	42.7	462	2 AH1053	probable exported
29	41	42.7	490	2 C70146	glutamate-tRNA lig

#### ALIGNMENTS

##### RESULT 1

C38128

epithelin/granulin precursor - mouse

N:Alternate names: acrogranin; PC-cell-derived growth factor.

C:Species: Mus musculus (house mouse)

C:Date: 10-Jul-1992 #sequence\_revision 10-Jul-1992 #text\_change 20-Aug-1999

C:Accession: C38128; S32503; I49468; A46705

R:Plowman, G.D.; Green, J.M.; Neubauer, M.G.; Buckley, S.D.; McDonald, V.L.; Todt

J. Biol. Chem. 267, 13073-13078, 1992

A:Title: The epithelin precursor encodes two proteins with opposing activities or

A:Reference number: A38128; MUID:92317004; PMID:1618805

A:Accession: C38128

A:Molecule type: mRNA

A:Residues: 1-589 <PI>

A:Cross-references: GB:X62321; NID:950851; PIDN:CAA44197.1; PID:950852

R:Baba, T.; Nemoto, H.; Watanabe, K.; Arai, Y.; Gerton, G.L.

FEBS Lett. 322, 89-94, 1993

A:Title: Exon/Intron organization of the gene encoding the mouse epithelin/granu.

A:Reference number: S32503; MUID:93245991; PMID:8482392

A:Accession: S32503

A:Molecule type: DNA

A:Residues: 18-349, 'L', 351-589 <BAB>

R:Baba, T.; Hoff, H.B.

Mol. Reprod. Dev. 34, 233-243, 1993

A:Title: Acrogranin, an acrosomal cysteine-rich glycoprotein, is the precursor of

A:Reference number: I48141; MUID:93228994; PMID:8471244

A:Accession: I49468

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-250, 'L', 252-253, 'V', 255-349, 'L', 351-401, 'SA', 404-589 <RES>

A:Cross-references: GB:M66736; NID:9191766; PIDN:AAA37191.1; PID:g191767

R:Zhou, J.; Gao, G.; Grabb, J.W.; Serrero, G.

J. Biol. Chem. 268, 10863-10869, 1993

A:Title: Purification of an autocrine growth factor homologous with mouse epithe.

A:Reference number: A46705; MUID:93266526; PMID:8496151

A:Accession: A46705

A>Status: preliminary

A:Molecule type: Protein

A:Residues: 18-19, 'X', 21-25, 'X', 27-29, 'XX', 32, 'XXX', 119-127, 152-154, 'DXK', 158-16;

C:Superfamily: granulin

Query Match 100.0%; Score 96; DB 2; Length 589;

Best Local Similarity 100.0%; Pred. No. 1.2e-07; Mismatches 0; Indels 0; Gaps 0;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVAPRRLPDPQILKSDT 19

|||||

DB 344 KKVAPRRLPDPQILKSDT 362

##### RESULT 2

B38128

RESULT 3  
 F81317  
 UDPglucose 4-epimerase (EC 5.1.3.2) Cj1131c [Imported] - Campylobacter jejuni (strain NC  
 C:Species: Campylobacter jejuni  
 C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 03-Jun-2002  
 C:Accession: F81317  
 R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling  
 C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell  
 Nature 403, 665-668, 2000  
 A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp  
 A:Reference number: AB1250; MUID:20150912; PMID:10688204  
 A:Accession: F81317  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-328 <PAR>  
 A:Cross-references: GB:AL11168; NID:G6968444; PIDN:CAB73386.1; PID:G696856  
 A:Experimental source: serotype O2, strain NCTC 11168  
 C:Genetics:  
 A:Gene: galE; Cj1131c  
 C:Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology  
 C:Keywords: isomerase  
  
 Query Match 50.0%; Score 48; DB 2; Length 328;  
 Best Local Similarity 55.6%; Pred. No. 4.2;

R; Bateman, A.; Belcourt, D.; Bennett, H.; Lazure, C.; Solomon, S. Biochem. Biophys. Res. Commun. 173, 1161-1168, 1990

A:Title: Granulins, a novel class of peptide from leukocytes.  
A:Reference number: A36698; MUID:91097544; PMID:2268320  
A:Accession: A36698  
A:Molecule type: protein  
A:Residues: 281-336 <BAT>  
A:Note: this protein was purified and characterized as granulins A  
A:Accession: B36698  
A:Molecule type: protein  
A:Residues: 206-218, 'H', 220-233 <BA2>  
A:Note: this protein was purified and characterized as granulins B  
A:Accession: C36698  
A:Molecule type: protein  
A:Residues: 364-367, 'X', 369-385, 'H', 387-396 <BA3>  
A:Note: this protein was purified and characterized as granulins C  
A:Accession: D36698  
A:Molecule type: protein  
A:Residues: 442-446, 'X', 456-458, 'DG' <BA4>  
R: Kardana, A.; Bagshawe, K.D.; Coles, B.; Read, D.; Taylor, M.  
Br. J. Cancer 67, 686-692, 1993  
A:Title: Characterisation of UGP and its relationship with beta-core fragment.  
A:Reference number: A56873; MUID:93229246; PMID:8471426  
A:Accession: A56873  
A:Molecule type: protein  
A:Residues: 281-283, 'X', 285-289, 'S', 291-295 <KAR>  
A:Experimental source: urine  
A:Note: sequence extracted from NCBI backbone (NCBIP:129524)  
C:Genetics:  
A:Gene: GDB:GRN  
A:Cross-references: GDB:136006; OMIM:138945  
A:Map position: 17pter-17qter  
A:Introns: 46/3; 88/3; 117/1; 154/3; 200/1; 236/3; 279/1; 311/3; 393/3; 471/3; 548/3  
C:Superfamily: granulins  
C:Keywords: glycoprotein; tandem repeat  
F:1-17/Domain: signal sequence #status predicted <SIG>  
F:18-593/Product: granulin #status predicted <BAT>  
F:18-593/Product: paragrulin #status predicted <PRO>  
F:18-44/Product: paragrulin #status experimental <PGR>  
F:58-113/Product: granulin G #status predicted <GRG>  
F:123-179/Product: granulin F #status predicted <GRF>  
F:206-261/Product: granulin B #status experimental <GRB>  
F:281-336/Product: granulin A #status experimental <GRA>  
F:364-417/Product: granulin C #status experimental <GRC>  
F:442-496/Product: granulin D #status predicted <GRD>  
F:518-573/Product: granulin E #status predicted <GRE>  
F:368/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 45.88; Score 44; DB 1; Length 593;  
Best Local Similarity 55.68; Pred. No. 38;  
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 KKVIAPRRLPDQILKSD 18  
DB 346 EKAPALSLPDQALKRD 363

RESULT 6  
T07099  
late embryogenesis (Lea)-like protein - tomato  
C:Species: Lycopersicon esculentum (tomato)  
C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 01-Dec-2000  
C:Accession: T07099; S49580  
R: van der Eycken, W.V.  
submitted to the EMBL Data Library, August 1997  
A:Reference number: Z15918  
A:Accession: T07099  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-174 <VAN>  
A:Cross-references: EMBL:Z46654; PIDN:CA86613.1  
A:Experimental source: strain Marmande; Meloidogyne incognita-induced root knot (gall)  
R: van der Eycken, W.; de Almeida Engler, J.; Inze, D.; van Montagu, M.; Gheysen, G.  
submitted to the EMBL Data Library, November 1994  
A:Description: A molecular study of Meloidogyne incognita-induced feeding sites.

A:Reference number: S49580  
A:Accession: S49580  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-138, 'RRV' <VA2>  
A:Cross-references: EMBL:Z46654

Query Match 44.88; Score 43; DB 2; Length 174  
Best Local Similarity 47.18; Pred. No. 13;  
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 2 KVIAPRLPDQILKSD 18  
DB 68 RVIASGRIPDPGSIKAN 84

RESULT 7  
T00150  
hypothetical protein 17 - Staphylococcus aureus phage phi PVL  
C:Species: Staphylococcus aureus phage phi PVL  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 11-May-2000  
C:Accession: T00150  
R: Kaneko, J.; Kimura, T.; Kawakami, Y.; Tomita, T.; Kamio, Y.  
Biosci. Biotechnol. Biochem. 61, 1960-1962, 1997  
A:Title: Pantone-Valentine leukocidin genes in a phage-like particle isolated from  
A:Reference number: Z14119; MUID:98067870; PMID:9404084  
A:Accession: T00150  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-223 <KAN>  
A:Cross-references: EMBL:AB009866; NID:dl204727; PIDN:BAA31890.1; PID:dl032851

Query Match 44.88; Score 43; DB 2; Length 223;  
Best Local Similarity 57.18; Pred. No. 18;  
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 6 PRLPDQILKSDT 19  
DB 141 PNNLPNSISNSDT 154

RESULT 8  
H87710  
hypothetical protein CC3722 [Imported] - Caulobacter crescentus  
C:Species: Caulobacter crescentus  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 10-May-2001  
C:Accession: H87710  
R: Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, J.C.; et al.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of Caulobacter crescentus.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: H87710  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-344 <STO>  
A:Cross-references: GB:AE005673; NID:gl3425492; PIDN:AAK25684.1; GSPDB:GN00148  
C:Genetics:  
C:Superfamily: phosphoglycerate dehydrogenase

Query Match 44.88; Score 43; DB 2; Length 344;  
Best Local Similarity 72.78; Pred. No. 29;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 KVIAPRLPD 12  
DB 23 KVIIVRLPD 33

RESULT 9  
S56142

calcium-binding protein precursor cnx1 - fission yeast (Schizosaccharomyces pombe)

N;Alternate names: calnexin homolog cnx1  
C:Species: Schizosaccharomyces pombe  
C:Date: 28-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 10-Dec-1999  
C:Accession: S56142; A56106; T38697  
R:Parlati, F.; Dignard, D.; Bergeron, J.J.M.; Thomas, D.Y.  
EMBL J. 14, 3064-3072, 1995  
A:Title: The calnexin homologue cnx1(+) in Schizosaccharomyces pombe, is an essential gene  
A:Reference number: S56142; MUID:95347333; PMID:7621821  
A:Accession: S56142  
A:Molecule type: DNA  
A:Residues: 1-560 <FAR>  
A:Cross-references: EMBL:M98799; NID:g437733; PIDN:AAA9757.1; PID:g437734  
J. J. Biol. Chem. 270, 4845-4853, 1995  
R:Jannatipour, M.; Rokeach, L.A.  
A:Title: The Schizosaccharomyces pombe homologue of the chaperone calnexin is essential  
A:Reference number: A56106; MUID:95181485; PMID:7876257  
A:Accession: A56106  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-560 <JAN>  
A:Cross-references: GB:U13389; NID:g532789; PIDN:AAA58631.1; PID:g532790  
A:Note: authors translated the codon GCA for residue 491 as Ser  
R:Brown, D.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, September 1997  
A:Reference number: Z21805  
A:Accession: T38697  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-560 <BRO>  
A:Cross-references: EMBL:Z99568; PIDN:CAB16741.1; GSPDB:GN00066; SPDB:SPAC3C7.11c  
A:Experimental source: strain 972h-; cosmid c3C7  
C:Genetics:  
A:Gene: SPDB:SPAC3C7.11c  
A:Map position: 1  
C:Superfamily: calnexin  
C:Keywords: calcium binding

Query Match 44.8%; Score 43; DB 2; Length 560;  
Best Local Similarity 50.0%; Pred. No. 52;  
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 APRRLPDPQILKSD 18  
||| :||| :| :  
Db 274 APRMPPDPVAVKPE 287

RESULT 10  
T01619  
hypothetical protein At2g18910 [imported] - Arabidopsis thaliana  
N;Alternate names: hypothetical protein F19f24.11  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 19-Feb-1999 #sequence\_revision 19-Feb-1999 #text\_change 02-Feb-2001  
C:Accession: T01619; B84570  
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul  
submitted to the EMBL Data Library, April 1998  
A:Description: Arabidopsis thaliana chromosome II BAC F19F24 genomic sequence.  
A:Reference number: Z14153  
A:Accession: T01619  
A>Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-753 <ROU>  
A:Cross-references: EMBL:AC003673; NID:g3004543; PID:g3004565  
A:Experimental source: cultivar Columbia  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: B84570  
A:Accession: B84570  
A>Status: preliminary

Query Match 44.8%; Score 43; DB 2; Length 753;  
Best Local Similarity 58.8%; Pred. No. 73;  
Matches 10; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 KVIAPRRRLPDPQILKS 17  
||| :||| :| :  
Db 665 KPWMAVPLPPPPQPKS 681

RESULT 11  
S74911  
hypothetical protein sll1950 - Synecocystis sp. (strain PCC 6803)  
C:Species: Synecocystis sp.  
A:Variety: PCC 6803  
C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999  
C:Accession: S74911  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,  
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.;  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecoch  
s.  
A:Reference number: S74322; MUID:97061201; PMID:8905231  
A:Accession: S74911  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-244 <KAN>  
A:Cross-references: EMBL:D90909; GB:AB001339; NID:g1652844; PIDN:BAAL7872.1; PID:d  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 43.8%; Score 42; DB 2; Length 244;  
Best Local Similarity 63.6%; Pred. No. 29;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 9 LPDPQILKSDT 19  
||| :||| :| :  
Db 148 LPQPOVIKIDT 158

RESULT 12  
D70987  
probable plcD protein - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
C:Accession: D70987  
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gort  
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroy  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete ge  
A:Reference number: A70500; MUID:98295987; PMID:9634230  
A:Accession: D70987  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-280 <COL>  
A:Cross-references: GB:Z95890; GB:AL123456; NID:g3242245; PIDN:CAB09340.1; PID:el30  
A:Experimental source: strain H37RV  
C:Genetics:  
A:Gene: plcD

Query Match 43.8%; Score 42; DB 2; Length 280;  
Best Local Similarity 70.0%; Pred. No. 34;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 6 PRRLPDPQIL 15



A:Reference number: I56399; MUID:85264808; PMID:2410621  
A:Accession: I73480  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-10 <LYN>  
A:Cross-references: GB:M28570; NID:g290476; PIDN:AAA24673.1; PID:g290478  
C:Genetics:  
A:Gene: thrA; thrA1; thrA2  
A:Map position: 0 min  
C:Function: <ASP>  
A:Description: as aspartate kinase, catalyzes the phosphorylation by ATP of aspartate.  
A:Pathway: aspartate metabolism  
A:Note: regulated allosterically by L-threonine  
C:Function: <HOM>  
A:Description: as homoserine dehydrogenase catalyzes the reduction by NADH of aspartate  
A:Pathway: glycine metabolism; serine metabolism; threonine biosynthesis  
C:Superfamily: thrA bifunctional enzyme; aspartate kinase homology; homoserine dehydroge  
C:Keywords: allosteric regulation; homotetramer; multifunctional enzyme; oxidoreductase;  
F:1-460/Domain: aspartate kinase homology <DKI>  
F:463-719/Domain: homoserine dehydrogenase homology <HSD>  
F:467-495/Region: beta-alpha-beta NAD(P) nucleotide-binding fold  
Query Match 43.8%; Score 42; DB 1; Length 820;  
Best Local Similarity 58.3%; Pred. NO. 1.2e+02;  
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
QY 6 PRLPDPOLKS 17  
Db 237 PRQVPDRLKS 248  
||:|:|:|:|

Search completed: July 8, 2003, 16:29:22  
Job time : 22.2436 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2003, 16:16:30 ; Search time 9.25641 Seconds  
(without alignments)  
85.136 Million cell updates/sec

Title: US-09-824-647-3

Perfect score: 96

Sequence: 1 KKVAPRRRLPQILKSDT 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	89	92.7	589	1 GRN_MOUSE	P28798 mus musculus
2	62	64.6	588	1 GRN_RAT	P23785 r granulin
3	46	47.9	824	1 AD08_HUMAN	P78325 homo sapien
4	44.5	46.4	256	1 RHLG_PSEAE	Q9RPT1 pseudomonas
5	44	45.8	593	1 GRN_HUMAN	P28799 h granulin
6	43	44.8	560	1 CALX_SCHPO	P36581 schizosacch
7	42	43.8	514	1 PHLD_MYCTU	Q9XB33 mycobacteri
8	42	43.8	795	1 PGN_HUMAN	Q9UQ90 homo sapien
9	42	43.8	820	1 AK1H_ECOLI	P00561 escherichia
10	42	43.8	867	1 POL_SRV2	P51517 simian retr
11	42	43.8	874	1 DNAB_MYCTU	P71715 mycobacteri
12	41.5	43.2	622	1 AMAL_PLAF8	P50492 plasmodium
13	41.5	43.2	622	1 AMAL_PLAFH	P22621 plasmodium
14	41.5	43.2	622	1 AMAL_PLAFH	P50491 plasmodium
15	41	42.7	490	1 SYE_BORBU	O51345 borrelia bu
16	40.5	42.2	643	1 SGT1_ARATH	Q91SM5 arabidopsis
17	40.5	42.2	1025	1 ADAL_YEAST	P38065 saccharomyc
18	40	41.7	241	1 OMPA_ESCBL	Q99124 escherichia
19	40	41.7	256	1 PGLG_CHLMU	Q9PKK7 chlamydia m
20	40	41.7	257	1 VGLG_BRVSU	O10686 bovine resp
21	40	41.7	293	1 CGD3_RAT	P48961 rattus norv
22	40	41.7	328	1 GALE_CORDI	P33119 corynebacte
23	40	41.7	330	1 UL16_HSVSA	Q01022 herpesvirus
24	40	41.7	405	1 VLI_HP26	Q9UYV6 pyrococcus
25	40	41.7	503	1 VLI_HP26	P36735 human papil
26	40	41.7	534	1 IF2P_SULAC	P95691 sulfolobus
27	40	41.7	819	1 AK1H_SERMA	P57725 serratia ma
28	40	41.7	830	1 GYRA_BUCAI	P57277 buchera ap
29	40	41.7	970	1 Y852_HUMAN	Q9Y629 homo sapien
30	39	40.6	126	1 PSAG_CHURE	P14224 chlamydomon
31	39	40.6	268	1 LPXH_RALSO	Q9Y081 raistonomye
32	39	40.6	329	1 GALE_STRLI	P13226 streptomyce
33	39	40.6	491	1 UNC3_CABEL	Q93705 caenorhabdi

34	39	40.6	726	1 PBPA_AQAE	O66874 aquifex aeo
35	39	40.6	739	1 UL47_HSVBC	P36338 bovine herp
36	38.5	40.1	152	1 LE14_SOYBN	P46519 glycine max
37	38.5	40.1	504	1 VLI_HP31	P17388 human papil
38	38.5	40.1	524	1 VLI_HP58	P26535 human papil
39	38.5	40.1	622	1 AMAL_PLAF8	P50489 plasmodium
40	38.5	40.1	622	1 AMAL_PLAF8	P50490 plasmodium
41	38	39.6	227	1 EXB2_SYNY3	P72944 synechocyst
42	38	39.6	291	1 CGD2_CHICK	P49706 gallus sapie
43	38	39.6	292	1 CGD3_HUMAN	P30281 homo sapien
44	38	39.6	292	1 CGD3_MOUSE	P30282 mus musculu
45	38	39.6	310	1 ATPG_SPIPL	P50006 spirulina p

#### ALIGNMENTS

RESULT 1  
GRN\_MOUSE  
ID GRN\_MOUSE STANDARD; PRT; 589 AA.  
AC P28798;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Granulins precursor (Acrogranin) [Contains: Granulin 1; Granulin 2; Granulin 3; Granulin 4; Granulin 5; Granulin 6; Granulin 7].  
GN GRN.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93245991; PubMed=8483292;  
RA Baba T., Nemoto H., Watanabe K., Arai Y., Gerton G.L.;  
RT "Exon/intron organization of the gene encoding the mouse  
epithelin/granulin precursor (acrogranin).";  
RT FEBS Lett. 322:89-94(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RX MEDLINE=32317004; PubMed=1618805;  
RA Plowman G.D., Green I.M., Neubauer M.G., Buckley S., McDonald V.L.,  
RT Todaro G.I., Shoyab M.;  
RT "The epithelin precursor encodes two proteins with opposing  
activities on epithelial cell growth.";  
RT J. Biol. Chem. 267:13073-13078(1992).  
CC -!- FUNCTION: GRANULINS HAVE POSSIBLE CYTOKINE-LIKE ACTIVITY. THEY MAY  
PLAY A ROLE IN INFLAMMATION, WOUND REPAIR, AND TISSUE REMODELLING.  
CC -!- TISSUE SPECIFICITY: UBIQUITOUS.  
CC -!- PTM: GRANULINS ARE DISULFIDE BRIDGED.  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; D16195; BAA03736.1; -;  
CC EMBL; M86736; AAA37191.1; -;  
CC EMBL; X62321; CAA44197.1; -;  
CC MGD; MGI:95832; Grn.  
CC InterPro; IPR000118; Granulin.  
CC Pfam; PF00396; granulin; 7.  
CC SMART; SM00277; GRAN; 7.  
CC PROSITE; PS00799; GRANULINS; 7.  
CC Cytokine; Repeat; Glycoprotein; Signal.  
CC SIGNAL 1 17  
CC CHAIN 18 589  
CC PEPTIDE 758 7113  
CC ACROGRANIN.  
CC GRANULIN 1.  
CC PEPTIDE 7122 7178  
CC GRANULIN 2.



```

CC CC -!- PTM: GRANULINS ARE DISULFIDE BRIDGED.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licensed@isb-sib.ch).
CC -----
CC EMBL; M97750; AAA16903.1; -.
CC EMBL; X62322; CAA44198.1; -.
CC PIR; A36199; A36199.
CC PIR; B36199; B36199.
CC PIR; E36698; E36698.
CC InterPro; IPR000118; Granulin..
CC Pfam; PF00396; granulin; 7.
CC SMART; SM00277; GRAN: 7.
CC PROSITE; PS00799; GRANULINS; 7.
KW Cytochrome; Repeat; Glycoprotein; Signal.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 588 ACROGRANIN.
FT PEPTIDE 58 113 GRANULIN 1.
FT PEPTIDE 122 178 GRANULIN 2.
FT PEPTIDE 204 259 GRANULIN 3.
FT PEPTIDE 278 334 GRANULIN 4.
FT PEPTIDE 361 413 GRANULIN 5.
FT PEPTIDE 418 492 GRANULIN 6.
FT PEPTIDE 512 567 GRANULIN 7.
FT CARBOHYD 38 38 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 372 372 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 525 525 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CONFLICT 201 201 S -> FP (IN REF. 2).
FT CONFLICT 307 308 TK -> SB (IN REF. 4).
FT CONFLICT 324 324 Q -> T (IN REF. 4).
FT CONFLICT 388 388 M -> I (IN REF. 2).
SQ SEQUENCE 588 AA; 63369 MW; 113D434FE099B31 CRC64;

Query Match 64.6%; Score 62; DB 1; Length 588;
Best Local Similarity 72.2%; Pred. No. 0.016;
Matches 13; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KKVIAPRRLPDPOILKSD 18
   ||| | ||||| |
Db 343 KKVTSLSLPDPQILKND 360

RESULT 3
AD08_HUMAN STANDARD; PRT; 824 AA.
ID AD08_HUMAN
AC P78325;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE ADAM 8 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE domain 8) (Cell surface antigen MS2) (CD156a antigen) (CD156).
GN ADAM8 OR MS2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_Taxid=9606;
[1]
RN SEQUENCE FROM N.A.
RP RC
RP TISSUE-Blood;
RX MEDLINE-97271556; PubMed-9126482;
RA Yoshizawa K., Higuchi Y., Kataoka M., Matsuura K., Yamamoto S.;
RT *CD156 (human ADAM8); expression, primary amino acid sequence, and
RT gene location.*
RL Genomics 41:56-62(1997).
CC -!- FUNCTION: POSSIBLE INVOLVEMENT IN EXTRAVASATION OF LEUKOCYTES.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: EXPRESSED ON NEUTROPHILS AND MONOCYTES.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
```

```

CC CC -1- SIMILARITY: CONTAINS 1 DISINTEGRIN DOMAIN.
CC CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD156 entry;
CC CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd156.htm"
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch)
CC CC -----
CC DR EMBL; D26579; BAA05626.1; -
CC DR HSP; P18619; IFVL.
CC DR MEROPS; M12.208; -
CC DR Genew; HGNC:215; ADAM8.
CC DR MIN; 602267; -
CC DR InterPro; IPR001762; Disintegrin.
CC DR InterPro; IPR000561; EGF-like.
CC DR InterPro; IPR002870; Pep.M12B.propep.
CC DR InterPro; IPR001590; Reprolysin.
CC DR InterPro; IPR000130; Zn.Mpeptidase.
CC DR Pfam; PF00200; disintegrin; 1.
CC DR Pfam; PF01421; Reprolysin; 1.
CC DR Pfam; PF01562; Pep.M12B.propep; 1.
CC DR ProDom; PD000664; Disintegrin; 1.
CC DR SMART; SM00050; DISIN; 1.
CC DR SMART; SM00181; EGF; 1.
CC DR PROSITE; PS00215; ADAM.MEPRO; 1.
CC DR PROSITE; PS00427; DISINTEGRIN.1; FALSE_NEG..
CC DR PROSITE; PS0214; DISINTEGRIN.2; 1.
CC DR PROSITE; PS01186; EGF.2; UNKNOWN.1.
CC DR PROSITE; PS00142; ZINC_PROTEASE; 1.
CC DR Hydrolase; Metalloprotease; zinc; Signal; Glycoprotein;
CC KW Transmembrane; Antigen.
CC FT SIGNAL 1 16 POTENTIAL.
CC FT CHAIN 17 824 ADAM 8.
CC FT DOMAIN 17 655 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 656 676 POTENTIAL.
CC FT DOMAIN 677 824 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 200 400 METALLOPROTEASE.
CC FT DOMAIN 408 494 DISINTEGRIN-LIKE.
CC FT METAL 334 334 ZINC (CATALYTIC) (PROBABLE).
CC FT ACT_SITE 335 335 BY SIMILARITY.
CC FT METAL 338 338 ZINC (CATALYTIC) (PROBABLE).
CC FT METAL 344 344 ZINC (CATALYTIC) (PROBABLE).
CC FT DISULFID 310 395 BY SIMILARITY.
CC FT CARBOHYD 67 67 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 91 91 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 436 436 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 612 612 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT SEQUENCE 824 AA; 88673 MW; 5DF8E05F30DF479E CRC64;
CC Query Match 47.9%; Score 46; DB 1; Length 824;
CC Best Local Similarity 43.8%; Pred. No. 9.8;
CC Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;
CC QY 2 KVIAPRRLPDPQILKS 17
CC DB 30 EVVLPRRLPGPRVRA 45
CC RESULT 4
CC RHLG_PSEAE STANDARD; PRT; 256 AA.
CC AC Q9RPT1;
CC DT 16-OCT-2001 (Rel. 40, Created)
CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Rhamnolipids biosynthesis 3-oxoacyl-[acyl-carrier protein] reductase
CC DE (EC 1.1.1.100) (3-ketoacyl-acyl carrier protein reductase).
CC GN RHLG OR PA3387.
CC OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OC NCBI_TaxID=287;
OC [1]
OC RP SEQUENCE FROM N.A.
OC RC STRAIN=ATCC 15692 / PA01;
OC RX MEDLINE=38389657; PubMed=3721281;
OC RA Campos-Garcia J., Caro A.D., Najera R., Miller-Mater R.M.,
OC Al-Tahhan R.A., Soberon-Chavez G.;
OC *The Pseudomonas aeruginosa rhlg gene encodes an NADPH-dependent beta-
OC ketoacyl reductase which is specifically involved in rhamnolipid
OC synthesis.*;
OC FT J. Bacteriol. 180:4442-4451(1998).
OC [2]
OC RP SEQUENCE FROM N.A.
OC RC STRAIN=ATCC 15692 / PA01;
OC RX MEDLINE=20437337; PubMed=10984043;
OC RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warriner P.,
OC Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
OC Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
OC Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
OC Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
OC Reizer J., Saler M.H., Hancock R.E.W., Lory S., Olson M.V.;
OC *Complete genome sequence of Pseudomonas aeruginosa PA01, an
OC opportunistic pathogen.*;
OC Nature 406:959-964(2000).
OC CC -1- FUNCTION: REQUIRED FOR THE SYNTHESIS OF THE BETA-HYDROXY ACID
OC MOIETY OF RHAMNOLIPIDS.
OC CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] +
OC NADPH(+) -> 3-oxoacyl-[acyl-carrier protein] + NADPH.
OC CC -1- PATHWAY: Rhamnolipids fatty acid moiety biosynthesis.
OC CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
OC (SDR) FAMILY.
OC CC -----
OC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
OC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
OC CC the European Bioinformatics Institute. There are no restrictions on its
OC CC use by non-profit institutions as long as its content is in no way
OC CC modified and this statement is not removed. Usage by and for commercial
OC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
OC CC or send an email to license@isb-sib.ch)
OC CC -----
OC DR EMBL; AF148964; AAD53514.1; -
OC DR EMBL; AE004760; AAG06775.1; -
OC DR HSP; P50162; 1AEL.
OC DR InterPro; IPR002198; ADH_short.
OC DR Pfam; PF00106; adh_short; 1.
OC DR PRINTS; PR00080; SDRFAMILY.
OC DR PROSITE; PS00061; ADH_SHORT; FALSE_NEG.
OC KW Oxidoreductase; NADP; Complete proteome.
OC FT NP_BIND 14 38 NADP (BY SIMILARITY).
OC FT ACT_SITE 160 160 BY SIMILARITY.
OC FT SEQUENCE 256 AA; 26830 MW; 3DB7B481F34C89A3 CRC64;
OC Query Match 46.4%; Score 44.5; DB 1; Length 256;
OC Best Local Similarity 42.3%; Pred. No. 4.7;
OC Matches 11; Conservative 3; Mismatches 3; Indels 9; Gaps 1;
OC QY 3 VIAPRRLP-----DPQILKSDT 19
OC DB 189 VIAPGRFFSRMTRHIANDPQALEADS 214
OC RESULT 5
OC GRN_HUMAN STANDARD; PRT; 593 AA.
OC AC P28799; P23781; P23782; P23783; P23784; Q9BWE7;
OC DT 01-NOV-1991 (Rel. 20, Created)
OC DT 01-DEC-1992 (Rel. 24, Last sequence update)
OC DT 15-JUN-2002 (Rel. 41, Last annotation update)
OC DE Granulins precursor (Acrogranin) [Contains: Paragranulin; Granulin 1
OC (Granulin G); Granulin 2 (Granulin F); Granulin 3 (Granulin B);
OC Granulin 4 (Granulin A); Granulin 5 (Granulin C); Granulin 6 (Granulin

```

D); Granulin 7 (Granulin E)].

GN GRN.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RC TISSUE=Bone marrow;

RX MEDLINE=92179253; PubMed=1542655;

RA Bhandari V., Palfrey R.G.E., Bateman A.;

RT "Isolation and sequence of the granulin precursor cDNA from human

RL bone marrow reveals tandem cysteine-rich granulin domains.";

RL Proc. Natl. Acad. Sci. U.S.A. 89:1715-1719(1992).

RN [2]

RP REVISIONS, SEQUENCE FROM N.A.

RX MEDLINE=93038704; PubMed=1417868;

RA Bhandari V., Bateman A.;

RT "Structure and chromosomal location of the human granulin gene.";

RL Biochem. Biophys. Res. Commun. 188:57-63(1992).

RN [3]

RP SEQUENCE FROM N.A.

RX TISSUE=Kidney;

RX MEDLINE=92317004; PubMed=1618805;

RA Plozman G.D., Green I.M., Neubauer M.G., Buckley S.D., McDonald V.L.,

RA Todaro G.J., Shoyab M.;

RT "The epithelin precursor encodes two proteins with opposing activities

RT on epithelial cell growth.";

RL J. Biol. Chem. 267:13073-13078(1992).

RN [4]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RA Yu W., Gibbs R.A.;

RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.

RN [5]

RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).

RC TISSUE=Cervix, and Lung;

RA Strausberg R.;

RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

RN [6]

RP SEQUENCE OF 206-233; 281-336; 364-396 AND 442-447.

RC TISSUE=Leukocyte;

RX MEDLINE=91097544; PubMed=2268320;

RA Bateman A., Belcourt D.R., Bennett H.P., Lazure C., Solomon S.;

RT "Granulins, a novel class of peptide from leukocytes.";

RL Biochem. Biophys. Res. Commun. 173:1161-1168(1990).

CC -!- FUNCTION: GRANULINS HAVE POSSIBLE CYTOKINE-LIKE ACTIVITY. THEY MAY

CC PLAY A ROLE IN INFLAMMATION, WOUND REPAIR, AND TISSUE REMODELING.

CC -!- FUNCTION: GRANULIN A PROMOTES PROLIFERATION OF THE EPITHELIAL

CC CELL LINE A431 IN CULTURE WHILE GRANULIN B ACTS AS AN ANTAGONIST

CC TO GRANULIN A, INHIBITING THE GROWTH.

CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; may be

CC produced by alternative splicing.

CC -!- TISSUE SPECIFICITY: IN MYELOGENOUS LEUKEMIC CELL LINES OF

CC PROMONOCYTIC, PROMYELOCYTIC, AND PROERYTHROID LINEAGE, IN

CC FIBROBLASTS, AND VERY STRONGLY IN EPITHELIAL CELL LINES. PRESENT

CC IN INFLAMMATORY CELLS AND BONE MARROW. HIGHEST LEVELS IN KIDNEY.

CC -!- PTM: GRANULINS ARE DISULFIDE BRIDGED.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; M75161; AAA58617.1; ALT\_SEQ.

DR EMBL; X62320; CAA44196.1; .

DR EMBL; AF055008; AAC09359.1; .

DR EMBL; BC000324; AAH00324.1; .

DR EMBL; BC010577; AAH10577.1; .

DR PIR; A38118; GYHU.

DR PIR; D36698; D36698.

DR PIR; JC1284; JC1284.

DR Genew; HGNC:4601; GRN.

DR MIM; 138945; .

DR InterPro; IPR000118; Granulin.

DR Pfam; PF00396; granulin; 7.

DR SMART; SM00277; GRAN; 7.

DR PROSITE; PS00799; GRANULINS; 7.

KW Cytokine; Repeat; Glycoprotein; Signal; Alternative splicing;

KW Polymorphism.

FT SIGNAL 1 17 POTENTIAL.

FT CHAIN 18 593 ACROGRANIN.

FT PEPTIDE 18 747 PARAGRANULIN.

FT PEPTIDE 18 7113 GRANULIN 1.

FT PEPTIDE 2123 2179 GRANULIN 2.

FT PEPTIDE 206 261 GRANULIN 3.

FT PEPTIDE 281 336 GRANULIN 4.

FT PEPTIDE 364 7417 GRANULIN 5.

FT PEPTIDE 442 7496 GRANULIN 6.

FT PEPTIDE 7518 7573 GRANULIN 7.

FT CARBOHYD 118 118 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 236 236 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 265 265 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 368 368 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 530 530 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT VARSPLIC 377 531 MISSING (IN ISOFORM 2).

FT VARIANT 454 454 G -> Q.

FT /FTID-VAR\_003445.

FT CONFLICT 219 219 S -> H (IN REF. 6).

FT CONFLICT 386 386 W -> H (IN REF. 6).

SQ SEQUENCE 593 AA; 63473 MW; 4E402BDB16DE2819 CRC64;

Query Match 45.88; Score 44; DB 1; Length 593;

Best Local Similarity 55.6%; Pred. No. 15;

Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 KKVIAPIRLPQILKSD 18

DB 346 EKAPAHLSLPDPQALKRD 363

RESULT 6

CALX SCHPO

ID CALX SCHPO STANDARD; PRT; 560 AA.

AC P36581;

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Calnexin homolog precursor

GN CAL1 OR CNX1 OR SPAC3C7.11C.

OS Schizosaccharomyces pombe (Fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.

OX NCBI\_TaxID=4896;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=95347333; PubMed=7621821;

RA Parlati F., Dignard D., Bergeron J.J.M., Thomas D.Y.;

RT "The calnexin homologue cnx1+ in Schizosaccharomycetes pombe, is an

RT essential gene which can be complemented by its soluble ER domain.";

RL EMBO J. 14:3064-3072(1995).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=95181485; PubMed=7876257;

RA Jannatipour M., Rokeach L.A.;

RT "The Schizosaccharomycetes pombe homologue of the chaperone calnexin is

RT essential for viability.";

RL J. Biol. Chem. 270:4845-4853(1995).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=972;

RX MEDLINE=21848401; PubMed=11859360;

RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
RA Mooney K., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,  
RA Wellens J., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
RA Cerruti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,  
RA "The genome sequence of Schizosaccharomyces pombe.";  
RL Nature 415:871-880(2002).  
CC -1- FUNCTION: CALCIUM-BINDING PROTEIN THAT INTERACTS WITH NEWLY  
CC SYNTHESIZED GLYCOPROTEINS IN THE ENDOPLASMIC RETICULUM. IT MAY ACT  
CC IN ASSISTING PROTEIN ASSEMBLY AND/OR IN THE RETENTION WITHIN THE  
CC ER OF UNASSEMBLED PROTEIN SUBUNITS. IT SEEMS TO PLAY A MAJOR ROLE  
CC IN THE QUALITY CONTROL APPARATUS OF THE ER BY THE RETENTION OF  
CC INCORRECTLY FOLDED PROTEINS.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Endoplasmic  
CC reticulum.  
CC -1- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M98799; AAA79757.1; -;  
CC EMBL; U13389; AAA68631.1; -;  
CC EMBL; Z99568; CAB16741.1; -;  
CC InterPro; IPR001580; Calreticulin.  
CC Pfam; PF00262; calreticulin; 1.  
CC PRINTS; PR00626; CALRETICULIN.  
CC ProDom; PD001866; Calreticulin; 1.  
CC PROSITE; PS00803; CALRETICULIN 1; 1.  
CC PROSITE; PS00804; CALRETICULIN 2; 1.  
CC PROSITE; PS00805; CALRETICULIN REPEAT; 2.  
CC Calcium-binding; Endoplasmic reticulum; Transmembrane; Signal;  
CC Repeat; Chaperone; Glycoprotein.  
CC SIGNAL 1 22 POTENTIAL.  
CC CHAIN 23 560 CALNEXIN HOMOLOG.  
CC DOMAIN 23 489 LUMENAL (POTENTIAL).  
CC TRANSMEM 490 512 POTENTIAL.  
CC DOMAIN 513 560 CYTOPLASMIC (POTENTIAL).  
CC DOMAIN 244 310 4 X APPROXIMATE REPEATS.  
CC REPEAT 244 255 1-1.  
CC REPEAT 261 272 1-2.  
CC REPEAT 280 291 1-3.  
CC REPEAT 299 310 1-4.  
CC DOMAIN 314 371 4 X APPROXIMATE REPEATS.  
CC REPEAT 314 324 2-1.  
CC REPEAT 333 343 2-2.  
CC REPEAT 347 357 2-3.  
CC REPEAT 361 371 2-4.  
CC CARBOHYD 418 N-LINKED (GLCNAC... ) (POTENTIAL).  
CC SEQUENCE 560 AA; 63466 MW; E3B20877333E9123 CRC64;

Query Match 44.8%; Score 43; DB 1; Length 560;  
Best Local Similarity 50.0%; Pred. No. 20;  
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
QY 5 APRRLPDQILKSD 18  
| | | | | : :  
Db 274 APRMIPDDAVKPE 287  
  
RESULT 7  
PHLD\_MYCTU  
ID PHLD\_MYCTU : STANDARD; PRT; 514 AA.  
AC Q9XB13; O06792;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Phospholipase C 4 precursor (EC 3.1.4.3).  
GN PLCD OR RV1755C OR MT1799 OR MTCY28.21C.  
OS Mycobacterium tuberculosis, and  
OS Mycobacterium bovis.  
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;  
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1773, 1765;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC SPECIES=M.tuberculosis; STRAIN=H37RV;  
RX MEDLINE=98295987; PubMed=9634230;  
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
RA Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. II, Tekala F.,  
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;  
RA "Deciphering the biology of Mycobacterium tuberculosis from the  
RA complete genome sequence.";  
RA Nature 393:537-544(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;  
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,  
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,  
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
RA Bishai W.;  
RA "Whole genome comparison of Mycobacterium tuberculosis clinical and  
RA laboratory strains.";  
RA Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC SPECIES=M.bovis; STRAIN=BCG / Pasteur;  
RX MEDLINE=99255698; PubMed=10320585;  
RA Gordon S.V., Brosch R., Billault A., Garnier T., Eiglmeyer K.,  
RA Cole S.T.;  
RA "Identification of variable regions in the genomes of tubercle bacilli  
RA using bacterial artificial chromosome arrays.";  
RA Mol. Microbiol. 32:643-655(1999).  
CC -1- CATALYTIC ACTIVITY: A phosphatidylcholine + H(2)O = 1,2-  
CC diacylglycerol + choline phosphate.  
CC -1- MISCELLANEOUS: POLYMORPHISM WAS DISCOVERED IN THE PHOSPHOLIPASE  
CC PLCD REGION.  
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PHOSPHOLIPASE C FAMILY.  
CC -1- CAUTION: The gene coding for this protein seems to be defective in  
CC strain H37RV where it is interrupted by the insertion of an IS6110  
CC element.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial

entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; 295890; CAB09340.1; ALT\_SEQ.  
CC EMBL; AE007040; NAK46072.1; -  
CC EMBL; Y18606; CAB44656.1; -  
DR TIGR; MT1799; -  
DR TubercuList; RV1755; -  
KW Hydrolyase; Signal; Complete proteome.  
FT SIGNAL 1 37 POTENTIAL.  
FT CHAIN 38 514 PHOSPHOLIPASE C 4.  
SQ SEQUENCE 514 AA; 55770 MW; 9D99A466D312BA30 CRC64;

Query Match 43.8%; Score 42; DB 1; Length 514;  
Best Local Similarity 70.0%; Pred. No. 26;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 6 PRRLPDPQIL 15  
I I I I I I I I  
Db 489 PYRPDPQIM 498

RESULT 8  
PGN\_HUMAN STANDARD; PRT; 795 AA.  
AC P9UQ90; 075756;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DE Paraplegin (EC 3.4.24.-) (Spastic paraplegia protein 7).  
GN PGN OR SPG7.  
OS Homo sapiens (Human).  
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
[1]  
SEQUENCE FROM N.A.  
RA Casari G., De Fusco M., Ciarmatori S., Zeviani M., Mora M.,  
RA Fernandez P., De Michele G., Filla A., Coccozza S., Marconi R.,  
RA Duerr A., Fontaine B., Ballabio A.;  
RT "Spastic paraplegia and OXPHOS impairment caused by mutations in  
paraplegin, a nuclear-encoded mitochondrial metalloprotease.";  
RL Cell 93:973-983(1998).  
[2]  
SEQUENCE FROM N.A.  
RX MEDLINE=9408245; PubMed=10480368;  
RA Sattasatian C., Whitmore S.A., Crawford J., Bilton R.L.,  
RA Cleton-Jansen A.-M., Sutherland G.R., Callen D.F.;  
RT "Genomic structure and expression analysis of the spastic paraplegia  
gene, SPG7.";  
RL Hum. Genet. 105:139-144(1999).  
CC -1- FUNCTION: PUTATIVE ATP-DEPENDENT PROTEASE.  
CC -1- SUBCELLULAR LOCATION: Mitochondrial.  
CC -1- TISSUE SPECIFICITY: UBIQUITOUS.  
CC -1- DISEASE: DEFECTS IN PGN ARE A CAUSE OF SPASTIC PARAPLEGIA-7  
(SPG7). A FORM OF AUTOSOMAL RECESSIVE HEREDITARY SPASTIC  
PARAPLEGIA (AR-HSP). HSP IS A GROUP OF INHERITED DEGENERATIVE  
SPINAL CORD DISORDERS CHARACTERIZED BY A SLOW, GRADUAL,  
PROGRESSIVE WEAKNESS AND SPASTICITY (STIFFNESS) OF THE LEGS.  
INITIAL SYMPTOMS MAY INCLUDE DIFFICULTY WITH BALANCE, WEAKNESS AND  
STIFFNESS IN THE LEGS, MUSCLE SPASMS, AND DRAGGING THE TOES WHEN  
WALKING. IN SOME FORMS OF THE DISORDER, BLADDER SYMPTOMS (SUCH AS  
INCONTINENCE) MAY APPEAR, OR THE WEAKNESS AND STIFFNESS MAY SPREAD  
TO OTHER PARTS OF THE BODY. RATE OF PROGRESSION AND THE SEVERITY  
OF SYMPTOMS IS QUITE VARIABLE.  
CC -1- SIMILARITY: BELONGS TO THE AAA FAMILY OF APPASES.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M41.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial  
entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; Y16610; CRA76314.1; -  
DR EMBL; AF080525; AAD28099.1; JOINED.  
DR EMBL; AF080511; AAD28099.1; JOINED.  
DR EMBL; AF080512; AAD28099.1; JOINED.  
DR EMBL; AF080513; AAD28099.1; JOINED.  
DR EMBL; AF080514; AAD28099.1; JOINED.  
DR EMBL; AF080515; AAD28099.1; JOINED.  
DR EMBL; AF080516; AAD28099.1; JOINED.  
DR EMBL; AF080517; AAD28099.1; JOINED.  
DR EMBL; AF080518; AAD28099.1; JOINED.  
DR EMBL; AF080519; AAD28099.1; JOINED.  
DR EMBL; AF080520; AAD28099.1; JOINED.  
DR EMBL; AF080521; AAD28099.1; JOINED.  
DR EMBL; AF080522; AAD28099.1; JOINED.  
DR EMBL; AF080523; AAD28099.1; JOINED.  
DR EMBL; AF080524; AAD28099.1; JOINED.  
DR HSSP; P32168; 1E94.  
DR MEROPS; M41.006;  
DR Genew; HGNC:11237; SPG7.  
DR MIN; 602783; -  
DR InterPro; IPR003593; AAA\_ATPase.  
DR InterPro; IPR003959; AAA\_ATPase\_Centr.  
DR InterPro; IPR000642; Peptidase\_M41.  
DR Pfam; PF00004; AAA; 1.  
DR Pfam; PF01434; Peptidase\_M41; 1.  
DR SMART; SM00382; AAA; 1.  
DR TIGRFAMs; TIGR01241; FtsH\_fam; 1.  
KW ATP-binding; Mitochondrion; Transmembrane; Hydrolase; Metalloprotease;  
Zinc.  
FT TRANSMEM 145 165 POTENTIAL.  
FT TRANSMEM 249 269 POTENTIAL.  
FT NP\_BIND 349 356 ATP (POTENTIAL).  
FT METAL 574 574 ZINC (CATALYTIC) (BY SIMILARITY).  
FT ACT\_SITE 575 575 BY SIMILARITY.  
FT METAL 578 578 ZINC (CATALYTIC) (BY SIMILARITY).  
FT CONFLICT 12 12 G -> R (IN REF. 1).  
SQ SEQUENCE 795 AA; 88135 MW; DA63283387517226 CRC64;

Query Match 43.8%; Score 42; DB 1; Length 795;  
Best Local Similarity 55.6%; Pred. No. 43;  
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;  
Oy 1 KKVIAPRRLPDPQILKSD 18  
I I I I I I I I I I  
Db 755 KKIAPQRWIDAQREKQD 772

RESULT 9  
AKIH\_ECOLI STANDARD; PRT; 820 AA.  
ID AKIH\_ECOLI  
AC P00561; Q47659;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Bifunctional aspartokinase/homoserine dehydrogenase I (AKI-HDI)  
[Includes: Aspartokinase I (EC 2.7.2.4); Homoserine dehydrogenase I  
(EC 1.1.1.3)].  
GN THRA OR THRA1 OR THRA2 OR B0002.  
OS Escherichia coli.  
CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
CC Escherichia.  
OX NCBI\_TaxID=562;  
[1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=8107247; PubMed=7003595;  
RA Katinka M., Cossart P., Sibilli L., Saint-Girons I., Chalvignac M.A.,  
le Bras G., Cohen G.N., Yaniv M.;  
RT "Nucleotide sequence of the thra gene of Escherichia coli.";  
RL Proc. Natl. Acad. Sci. U.S.A. 77:5730-5733(1980).

[2] SEQUENCE FROM N.A.  
 RP STRAIN-K12;  
 RC MEDLINE-92334977; PubMed-1630901;  
 RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,  
 RA Isono K., Mizobuchi K., Nakata A.;  
 RT "Systematic sequencing of the Escherichia coli genome: analysis of  
 RT the 0-2.4 min region.";  
 RL Nucleic Acids Res. 20:3305-3308(1992).  
 [3] SEQUENCE FROM N.A.  
 RP STRAIN-K12 / MG1655;  
 RC MEDLINE-95334362; PubMed-7610040;  
 RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,  
 RA Blattner F.R.;  
 RT "Analysis of the Escherichia coli genome VI: DNA sequence of the  
 RT region from 92.8 through 100 minutes.";  
 RL Nucleic Acids Res. 23:2105-2119(1995).  
 [4] SEQUENCE FROM N.A.  
 RP MEDLINE-80077291; PubMed-390305;  
 RA Cossart P., Katinka M., Yaniv M.;  
 RT "Construction and expression of a hybrid plasmid containing the  
 RT Escherichia coli thrA and thrB genes.";  
 RL Mol. Gen. Genet. 175:39-44(1979).  
 [5] SEQUENCE OF 51-129.  
 RP MEDLINE-80043179; PubMed-387092;  
 RA Sibilli L., le Bras G., Cossart P., Chalvignac M.A., le Bras G.,  
 RA Briley P.A., Cohen G.N.;  
 RT "The primary structure of Escherichia coli K 12 aspartokinase I -  
 RT homoserine dehydrogenase I: sequence of cyanogen bromide peptide CB  
 RT 3.";  
 RL Biochimie 61:733-739(1979).  
 [6] REVISION TO 11.  
 RP MEDLINE-83135751; PubMed-6298218;  
 RA Cossart P.;  
 RL Unpublished results, cited by:  
 RL Zakin M.M., Duchange N., Ferrara P., Cohen G.N.;  
 RL J. Biol. Chem. 258:3028-3031(1983).  
 CC -|- CATALYTIC ACTIVITY: L-homoserine + NAD(P)(+) -> L-aspartate 4-  
 CC semialdehyde + NAD(P)H.  
 CC -|- CATALYTIC ACTIVITY: ATP + L-aspartate -> ADP + 4-phospho-L-  
 CC aspartate.  
 CC -|- ENZYME REGULATION: THE ENZYME ACTIVITIES ARE REGULATED  
 CC ALLOSTERICALLY BY L-THREONINE.  
 CC -|- PATHWAY: CATALYZES 2 NONCONSECUTIVE REACTIONS IN THE COMMON  
 CC BIOSYNTHETIC PATHWAY LEADING FROM ASP TO DIAMINOPIMELATE AND LYS,  
 CC TO MET, AND TO THR AND ILE.  
 CC -|- SUBUNIT: HOMOTETRAMER.  
 CC -|- MISCELLANEOUS: ASPARTOKINASE II-HOMOSERINE DEHYDROGENASE II AND  
 CC ASPARTOKINASE III ALSO CATALYZE THE SAME REACTION(S).  
 CC -|- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE  
 CC ASPARTOKINASE FAMILY.  
 CC -|- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE  
 CC HOMOSERINE DEHYDROGENASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; V00361; CAA23660.1; ALT\_SEQ.  
 DR EMBL; J01706; AAA83914.1;  
 DR EMBL; X58872; CAA48734.1;  
 DR EMBL; D10483; BAA01286.1;  
 DR EMBL; U14003; AAA97301.1;  
 DR EMBL; AE000111; AAC73113.1;  
 DR EMBL; V00360; CAA23659.1;

DR EMBL; M28570; AAA24673.1;  
 DR EMBL; M10644; AAA24671.1;  
 DR PIR; A00671; DEECK.  
 DR PIR; S40531; S40531.  
 DR EcoGene; EG10998; thrA.  
 DR InterPro; IPR002912; ACT.  
 DR InterPro; IPR001048; Aspartate\_kinase.  
 DR InterPro; IPR001341; Aspartate\_kinase.  
 DR InterPro; IPR001342; Homoserine\_dh.  
 DR InterPro; IPR005106; NAD\_binding\_3.  
 DR Pfam; PF00696; aakinas; 1.  
 DR Pfam; PF00742; Homoserine\_dh; 1.  
 DR Pfam; PF01842; ACT; 2.  
 DR Pfam; PF03447; NAD\_binding\_3; 1.  
 DR TIGRFAMS; TIGR00657; asp\_kinases; 1.  
 DR PROSITE; PS00324; ASPARTOKINASE; 1.  
 DR PROSITE; PS01042; HOMOSERINE DEHYDROGENASE; 1.  
 KW Transferase; Kinase; Oxidoreductase; Threonine biosynthesis; NADP;  
 KW Allosteric enzyme; Multifunctional enzyme; Complete proteome.  
 FT DOMAIN 1 249 ASPARTOKINASE I.  
 FT DOMAIN 250 470 INTERFACE.  
 FT NP\_BIND 471 820 HOMOSERINE DEHYDROGENASE I.  
 FT NP\_BIND 471 478 NADP (POTENTIAL).  
 FT CONFLICT 113 113 Q -> E (IN REF. 5).  
 FT CONFLICT 230 230 N -> D (IN REF. 3).  
 FT CONFLICT 375 375 L -> Q (IN REF. 3).  
 FT CONFLICT 393 393 A -> T (IN REF. 3).  
 FT CONFLICT 406 406 L -> M (IN REF. 3).  
 FT CONFLICT 553 553 N -> D (IN REF. 3 AND 4).  
 FT CONFLICT 587 588 DY -> IT (IN REF. 4).  
 FT CONFLICT 607 607 I -> T (IN REF. 3).  
 FT CONFLICT 658 658 R -> T (IN REF. 3).  
 SQ SEQUENCE 820 AA; 89122 MW; E4DE917A5703F6E7 CRC64;

Query Match 43.8%; Score 42; DB 1; Length 820;  
 Best Local Similarity 58.3%; Pred. No. 44;  
 Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 6 PRRLPDPQILKS 17  
 ||:|:|:|:|  
 DB 237 PRQVPDARLLKS 248

RESULT 10  
 POL\_SRV2 STANDARD; PRT; 867 AA.  
 ID POL\_SRV2  
 AC P51517;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE POL polyprotein [Contains: Reverse transcriptase (EC 2.7.7.49);  
 DE Endonuclease].  
 GN POL.  
 OS Simian retrovirus SRV-2.  
 OC Viruses; Retrovirdae; Retroviridae.  
 OX NCBI\_TaxID=39068;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-87151131; PubMed-2435057;  
 RA Thayer R.M., Power M.D., Bryant M.L., Gardner M.B., Barr P.J.,  
 RA Luciw P.A.;  
 RT "Sequence relationships of type D retroviruses which cause simian  
 RT acquired immunodeficiency syndrome.";  
 RL Virology 157:317-329(1987).  
 CC -|- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate  
 CC + [DNA](N).  
 CC -|- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))

CC EMBL; M16605; AAA47562.1; -  
 CC HSSP; P03355; LMML  
 DR InterPro; IPR001037; Integrase\_C  
 DR InterPro; IPR003308; Integrase\_Zn  
 DR InterPro; IPR002156; RNaseH  
 DR InterPro; IPR000477; RVse  
 DR InterPro; IPR001584; Rve  
 DR Pfam; PF00075; rnaseh; 1  
 DR Pfam; PF00078; rvt; 1  
 DR Pfam; PF00552; integrase; 1  
 DR Pfam; PF00665; rve; 1  
 DR Pfam; PF02022; Integrase\_Zn; 1  
 KW Transferase; RNA-directed DNA polymerase; Hydrolase; Nuclease;  
 KW Endonuclease; Polyprotein  
 SQ SEQUENCE 867 AA; 98038 MW; F64227F7365F3659 CRC64;

Query Match 43.8%; Score 42; DB 1; Length 867;  
 Best Local Similarity 50.0%; Pred. No. 47;  
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 VIAPRRRLPDPQILKSD 18  
 :|||: || |||  
 Db 14 ILAPQYADPIITKSD 29

RESULT 11  
 DNAB\_MYCTU STANDARD; PRT; 874 AA.  
 AC P71715;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 13-DEC-1998 (Rel. 37, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Replicative DNA helicase (EC 3.6.1.-) [Contains: Endonuclease PI-  
 DE MtuHIP (EC 3.1.-) (Mtu dnab intein)]  
 GN DNAB OR RV0058 OR MT0064 OR MTCV21D4.21.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Actinobacteriia; Actinobacteria (class); Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_Taxid=1773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H37Rv;  
 RX MEDLINE=98295987; PubMed=9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Sulston J.E., Taylor K., Whitehead S., Squares S., Squares R.,  
 RA "Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence."  
 RL Nature 393:537-544(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CDC 1551 / Oshkosh;  
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,  
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,  
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
 RA Bishai W.;  
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
 RT laboratory strains."  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: PARTICIPATES IN INITIATION AND ELONGATION DURING  
 CC CHROMOSOME REPLICATION; IT EXHIBITS DNA-DEPENDENT ATPASE ACTIVITY  
 CC (BY SIMILARITY).  
 CC -!- FUNCTION: THE INTEIN IS AN ENDONUCLEASE (POTENTIAL).  
 CC -!- PTM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES

CC A POST-TRANSLATIONAL EXCISION OF THE INTERVENING REGION (INTEIN)  
 CC FOLLOWED BY PEPTIDE LIGATION.  
 CC -!- SIMILARITY: BELONGS TO THE HELICASE FAMILY. DNAB SUBFAMILY.  
 CC -!- SIMILARITY: IN THE INTEIN SECTION; BELONGS TO THE HOMING  
 CC ENDONUCLEASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))

DR EMBL; Z80775; CAB02519.1;  
 DR EMBL; AE006918; AAK44286.1; -  
 DR HSSP; P03005; LJWE.  
 DR REBASE; 4230; PI-MtuHIP.  
 DR TIGR; MT0064;  
 DR Tuberculist; Rv0058;  
 DR InterPro; IPR001198; Dnab\_helicase.  
 DR InterPro; IPR003586; Hedgehog\_hintc.  
 DR InterPro; IPR003587; Hedgehog\_hintn.  
 DR InterPro; IPR002203; Intein.  
 DR InterPro; IPR004042; Intein\_endonuc.  
 DR Pfam; PF00772; Dnab; 1  
 DR ProDom; PD000229; RecA; 1.  
 DR ProDom; PD002663; Dnab\_helicase; 1.  
 DR SMART; SM00305; Hintc; 1.  
 DR SMART; SM00306; Hintn; 1.  
 DR TIGRFAMS; TIGR00665; Dnab; 1.  
 DR PROSITE; PS50818; INTEIN\_CTER; 1.  
 DR PROSITE; PS50819; INTEIN\_NTER; 1.  
 DR PROSITE; PS50817; INTEIN\_NTER; 1.  
 KW DNA replication; DNA-binding; ATP-binding; Hydrolase; Helicase;  
 KW Primosome; Protein splicing; Autocatalytic cleavage; Nuclease;  
 KW Endonuclease; Intron homing; Complete proteome.  
 FT CHAIN 1 399 REPLICATIVE DNA HELICASE, 1ST PART.  
 FT CHAIN 400 815 ENDONUCLEASE PI-MTUHIP.  
 FT CHAIN 816 874 REPLICATIVE DNA HELICASE, 2ND PART.  
 FT NP\_BIND 226 233 ATP (POTENTIAL).  
 FT CONFLICT 552 R -> G (IN REF. 2).  
 SQ SEQUENCE 874 AA; 96916 MW; 894155A86DCB9D70 CRC64;

Query Match 43.8%; Score 42; DB 1; Length 874;  
 Best Local Similarity 54.5%; Pred. No. 47;  
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 KVIAPRRLPDP 12  
 :|||:|:  
 Db 491 RIAAPRRVPEP 501

RESULT 12

AMAL\_PLAF8 STANDARD; PRT; 622 AA.  
 AC P50492;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Apical membrane antigen 1 precursor (Merozoite surface antigen).  
 GN AMA-1 OR PF83.  
 OS Plasmodium falciparum (isolate 7G8).  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_Taxid=57286;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91101665; PubMed=2270110;  
 RA Thomas A.W., Waters A.P., Carr D.;  
 RT "Analysis of variation in PF83, an erythrocytic merozoite vaccine  
 RT candidate antigen of Plasmodium falciparum."  
 RL Mol. Biochem. Parasitol. 42:285-287(1990).  
 CC -!- FUNCTION: INVOLVED IN PARASITE INVASION OF ERYTHROCYTES.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- SIMILARITY: STRONG TO AMA-1 FROM P.CHABAUDI AND P.FRAGILE, AND TO  
CC PK66 FROM P.KNOWLEDGE.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M58548; AAA29721.1; -  
CC InterPro: IPR003298; Apmem\_Ag1.  
CC Pfam; PF02430; AMA-1; 1.  
CC PRINTS; PR01361; MEROZOITESA.  
CC Malaria; Signal; Transmembrane; Antigen; Glycoprotein.  
CC SIGNAL 1 24 POTENTIAL.  
CC CHAIN 25 622 APICAL MEMBRANE ANTIGEN 1.  
CC DOMAIN 25 546 EXTRACELLULAR (POTENTIAL).  
CC TRANSMEM 547 567 POTENTIAL.  
CC DOMAIN 568 622 CYTOPLASMIC (POTENTIAL).  
CC CARBOHYD 162 162 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC CARBOHYD 286 286 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC CARBOHYD 371 371 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC CARBOHYD 421 421 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC CARBOHYD 422 422 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC CARBOHYD 499 499 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC SEQUENCE 622 AA; 71990 MW; 1C9C8715D8E2915F CRC64;  
Query Match 43.2%; Score 41.5; DB 1; Length 622;  
Best Local Similarity 36.7%; Pred. No. 39;  
Matches 11; Conservative 5; Mismatches 3; Indels 11; Gaps 2;  
QY 1 KVIAPR-----RLP-DPQILKSDT 19  
DB 472 KTIAPRIFTSDDKSLKCPDPEIVSNST 501  
||:||||| : |||: :  
RESULT 13  
AMAL\_PLAFF STANDARD; PRT; 622 AA.  
AC P22621;  
DT 01-AUG-1991 (Rel. 19, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Apical membrane antigen 1 precursor (Merozoite surface antigen).  
GN AMA-1 OR PF83.  
OS Plasmodium falciparum (isolate FC27 / Papua New Guinea).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=5837;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89384584; PubMed=2701947;  
RA Peterson M.G., Marshall V.M., Smythe J.A., Crewther P.E., Lew A.,  
RA Silva R.F., Kemp D.J.;  
RT "Integral membrane protein located in the apical complex of  
RT Plasmodium falciparum".  
RL Mol. Cell. Biol. 9:3151-3154(1989).  
CC -1- FUNCTION: INVOLVED IN PARASITE INVASION OF ERYTHROCYTES.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- SIMILARITY: STRONG TO AMA-1 FROM P.CHABAUDI AND P.FRAGILE, AND TO  
CC PK66 FROM P.KNOWLEDGE.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M27133; AAA29475.1; -

DR EMBL; M27957; AAA29476.1; -  
DR EMBL; A08267; CAA00764.1; -  
DR PIR; A32499; A32499.  
DR InterPro: IPR003298; Apmem\_Ag1.  
DR Pfam; PF02430; AMA-1; 1.  
DR PRINTS; PR01361; MEROZOITESA.  
KW Malaria; Signal; Transmembrane; Antigen; Glycoprotein.  
FT SIGNAL 1 24 POTENTIAL.  
FT CHAIN 25 622 APICAL MEMBRANE ANTIGEN 1.  
FT DOMAIN 25 546 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 547 567 POTENTIAL.  
FT DOMAIN 568 622 CYTOPLASMIC (POTENTIAL).  
FT CARBOHYD 162 162 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 286 286 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 371 371 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 421 421 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 422 422 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 499 499 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CONFLICT 308 308 Q -> E (IN REF. 1; AAA29476).  
FT CONFLICT 332 332 I -> N (IN REF. 1; AAA29476).  
FT CONFLICT 407 407 Q -> H (IN REF. 1; AAA29476).  
FT CONFLICT 439 439 H -> N (IN REF. 1; AAA29476).  
FT CONFLICT 496 496 I -> M (IN REF. 1; AAA29476).  
FT CONFLICT 503 503 N -> R (IN REF. 1; AAA29476).  
SQ SEQUENCE 622 AA; 72009 MW; 7D41335E249FA18F CRC64;  
Query Match 43.2%; Score 41.5; DB 1; Length 622;  
Best Local Similarity 36.7%; Pred. No. 39;  
Matches 11; Conservative 5; Mismatches 3; Indels 11; Gaps 2;  
QY 1 KVIAPR-----RLP-DPQILKSDT 19  
DB 472 KTIAPRIFTSDDKSLKCPDPEIVSNST 501  
||:||||| : |||: :  
RESULT 14  
AMAL\_PLAFF STANDARD; PRT; 622 AA.  
AC P50491;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Apical membrane antigen 1 precursor (Merozoite surface antigen).  
GN AMA-1 OR PF83.  
OS Plasmodium falciparum (isolate thtn / Thailand).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=70151;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91101665; PubMed=22701110;  
RA Thomas A.W., Waters A.P., Carr D.;  
RT "Analysis of variation in pf83, an erythrocytic merozoite vaccine  
RT candidate antigen of Plasmodium falciparum".  
RL Mol. Biochem. Parasitol. 42:285-287(1990).  
CC -1- FUNCTION: INVOLVED IN PARASITE INVASION OF ERYTHROCYTES.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- SIMILARITY: STRONG TO AMA-1 FROM P.CHABAUDI AND P.FRAGILE, AND TO  
CC PK66 FROM P.KNOWLEDGE.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M58547; AAA29720.1; -  
CC InterPro: IPR003298; Apmem\_Ag1.  
CC Pfam; PF02430; AMA-1; 1.  
CC PRINTS; PR01361; MEROZOITESA.  
KW Malaria; Signal; Transmembrane; Antigen; Glycoprotein.  
FT SIGNAL 1 24 POTENTIAL.



FT CHAIN 25 622 APICAL MEMBRANE ANTIGEN 1.  
FT DOMAIN 25 546 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 547 567 POTENTIAL.  
FT DOMAIN 568 622 CYTOPLASMIC (POTENTIAL).  
FT CARBOHYD 286 286 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 371 371 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 421 421 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 422 422 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 499 499 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 622 AA; 71989 MW; 1FDFA53593C94CC5 CRC64;

Query Match 43.2%; Score 41.5; DB 1; Length 622;  
Best Local Similarity 36.7%; Pred. No. 39;  
Matches 11; Conservative 5; Mismatches 3; Indels 11; Gaps 2;

QY 1 KVIAPR-----RLP-DPQILKSDT 19  
||:||||| :|||:|:|  
Db 472 KRIAPRIFISDKSLKPCDPFIVSNST 501

RESULT 15  
ID SVE\_BORBU STANDARD; PRT; 490 AA.  
AC 051345;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Glutamyl-tRNA synthetase (EC 6.1.1.17) (Glutamate--tRNA ligase) (GLURS).  
GN GLTX OR B03072.  
OS Borrelia burgdorferi (Lyme disease spirochete).  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxID=139;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 35210 / B31;  
RX MEDLINE=98065943; PubMed=9403685;  
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,  
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,  
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D., Hanson M.,  
RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,  
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,  
RA Utterback T., Wathley L., McDonald L., Artiach P., Bowman J.,  
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,  
RA Smith H.O., Venter J.C.;  
RT "Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi."  
RL Nature 390:580-586(1997).  
CC -|- CATALYTIC ACTIVITY: ATP + L-glutamate + tRNA(Glu) = AMP +  
CC diphosphate + L-glutamyl-tRNA(Glu).  
CC -|- SUBUNIT: MONOMER (BY SIMILARITY).  
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -|- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL: AE001142; AAC66742.1; -  
CC HSP: P27000; 1GLN.  
CC TIGR: B03072; -  
CC InterPro: IPR004527; GLTX\_bact.  
CC InterPro: IPR000924; Glu\_tRNA-synt\_1c.  
CC InterPro: IPR001412; tRNA-synt\_1.  
CC Pfam: PF00749; tRNA-synt\_1c; 1.  
CC PRINTS: PR00987; TRNASYNTHGLU  
CC TIGRFAMs: TIGR00464; gltx\_bact; 1.  
CC PROSITE: PS00178; AA\_TRNA\_LIGASE\_I; FALSE\_NEG.  
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;

KW Complete proteome.  
FT SITE 9 19 "HIGH" REGION.  
FT SITE 251 255 "KMSKS" REGION.  
FT BINDING 254 254 ATP (BY SIMILARITY).  
SQ SEQUENCE 490 AA; 56751 MW; E7ED6AD2474934F4 CRC64;

Query Match 42.7%; Score 41; DB 1; Length 490;  
Best Local Similarity 88.9%; Pred. No. 36;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 10 DPQILKSD 18  
||| |||||  
Db 182 DPQVILKSD 190

Search completed: July 8, 2003, 16:24:59  
Job time : 11.2564 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2003, 16:17:10 ; Search time 40.9231 Seconds  
(without alignments)  
95.665 Million cell updates/sec

Title: US-09-824-647-3

Perfect score: 96

Sequence: 1 KKVAPRRLPDPQILKSDT 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	89	92.7	589	11 Q9D2V3	Q9d2v3 mus musculus
2	48	50.0	328	2 O86149	O86149 campylobact
3	48	50.0	328	16 Q9PNG3	Q9png3 campylobact
4	48	50.0	330	2 Q9K5D5	Q9k5d5 campylobact
5	45	46.9	755	5 Q9N4K2	Q9n4k2 caenorhabdi
6	44.5	46.4	482	12 Q8QVC6	Q8qvc6 colorado ti
7	44.5	46.4	684	12 Q9ENK9	Q9enk9 colorado ti
8	44	45.8	413	4 Q9H8S1	Q9h8s1 homo sapien
9	44	45.8	463	4 Q13648	Q13648 homo sapien
10	44	45.8	466	4 Q9HBT9	Q9hbt9 homo sapien
11	44	45.8	1300	17 Q8TNC8	Q8tnc8 methanosarc
12	43	44.8	131	10 Q64621	Q64621 arabidopsis
13	43	44.8	164	11 Q99P28	Q99p28 mus musculu
14	43	44.8	174	10 Q40159	Q40159 lycopersico
15	43	44.8	223	9 O80056	O80056 staphylococ
16	43	44.8	255	2 Q9F0G6	Q9f0g6 pseudomonas

17	43	44.8	271	11 Q9D1F2	Q9d1f2 mus musculu
18	43	44.8	344	16 Q9A244	Q9a244 caulobacter
19	43	44.8	596	4 Q9H1R3	Q9h1r3 homo sapien
20	43	44.8	692	5 Q9VWD0	Q9vwd0 drosophila
21	43	44.8	851	5 Q8SQU0	Q8squ0 encephalito
22	43	44.8	1550	9 Q8SDK4	Q8sdk4 staphylococ
23	43	44.8	1789	5 Q8T2E3	Q8t2e3 dictyostell
24	42	43.8	124	5 Q9SRN4	Q9srn4 drosophila
25	42	43.8	244	16 P73818	P73818 synecocyst
26	42	43.8	298	5 Q9W045	Q9w045 drosophila
27	42	43.8	305	17 Q97CD6	Q97cd6 thermoplasm
28	42	43.8	465	10 Q9FW59	Q9fw59 oryza sativ
29	42	43.8	505	12 Q9YUR2	Q9yur2 turkey aden
30	42	43.8	556	1 Q53641	Q53641 sulfolobus
31	42	43.8	566	16 P71838	P71838 mycobacteri
32	42	43.8	627	6 O62677	O62677 salmirci sci
33	42	43.8	632	6 O62678	O62678 actus trivi
34	42	43.8	644	5 Q9N814	Q9n814 trypanosoma
35	42	43.8	780	5 Q9V4N9	Q9v4n9 drosophila
36	42	43.8	780	5 Q9N812	Q9n812 trypanosoma
37	42	43.8	820	2 Q8RMX0	Q8rmx0 escherichia
38	42	43.8	820	16 Q8ZS19	Q8zsl9 salmonella
39	42	43.8	820	16 Q8ZSR7	Q8zsr7 salmonella
40	42	43.8	820	16 Q8XA84	Q8xa84 escherichia
41	42	43.8	867	15 Q9WQ04	Q9wq04 simian retr
42	42	43.8	867	15 Q9WQ00	Q9wq00 simian retr
43	42	43.8	1650	5 Q26816	Q26816 trypanosoma
44	42	43.8	1657	5 Q8WPT4	Q8wpt4 trypanosoma
45	42	43.8	1657	5 Q8WPT3	Q8wpt3 trypanosoma

ALIGNMENTS

RESULT 1

Q9D2V3	PRELIMINARY;	PRT;	589 AA.
ID	Q9D2V3		
AC	Q9D2V3;		
DT	01-JUN-2001 (TREMBlrel. 17, Created)		
DT	01-JUN-2001 (TREMBlrel. 17, Last sequence update)		
DT	01-JUN-2002 (TREMBlrel. 21, Last annotation update)		
DE	Adult male kidney cDNA, RIKEN full-length enriched library,		
DE	clone:0610012H06, full insert sequence.		
GN	GRN		
OS	Mus musculus (Mouse)		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=KIDNEY;		
RX	MEDLINE=21085660; PubMed=11217851;		
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,		
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,		
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,		
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,		
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,		
RA	Fleischmann W., Gaasterland T., Glissi C., King B., Kochiwa H.,		
RA	Kuehl P., Lewis S., Matsuo Y., Nikolaic I., Pesole G., Queckenbush J.,		
RA	Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,		
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,		
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,		
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,		
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,		
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,		
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,		
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,		
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Willing L.,		
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,		
RA	Hayashizaki Y.;		
FT	*Functional annotation of a full-length mouse cDNA collection.*;		
RL	Nature 409:685-690(2001).		
DR	EMBL; AK018744; BAB31384.1; -.		

```
DR MGD; MGI:95832; Grn.
DR InterPro; IPR000118; Granulin.
DR InterPro; IPR001211; PhospholipaseA2.
DR Pfam; PF00396; granulin; 7.
DR SMART; SM00277; GRAN; 7.
DR PROSITE; PS00799; GRANULINS; 7.
DR PROSITE; PS00118; PA2_HIS; UNKNOWN_1.
SQ SEQUENCE 589 AA; 63405 MW; 1DE8229CA13CA292 CRC64;

Query Match          92.78; Score 89; DB 11; Length 589;
Best Local Similarity 94.78; Pred. No. 2.2e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 KKVIAAPRLPPDQILKSDT 19
    ||||| ||||| ||||| |||||
Db 344 KKVIAAPRLPPDQILKSDT 362

RESULT 2
ID O86149 PRELIMINARY; PRT; 328 AA.
AC O86149;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE GALE protein.
GN GALE.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=81116;
RX MEDLINE=98386495; PubMed=9720026;
RA Fry B.N., Korolik V., ten Brinke J.A., Pennings M.T.T., Zalm R.,
RA Tunis B.J.J., Coloe P.J., van der Zeijst B.A.M.;
RT "The lipopolysaccharide biosynthesis locus of Campylobacter jejuni
RT 81116."
RL Microbiology 144:2049-2061(1998).
DR EMBL; Y11648; CAA72350.1; -.
DR HSSP; P09147; IXEL.
DR InterPro; IPR001509; Epimerase_Dh.
DR Pfam; PF01370; Epimerase; 1.
DR TIGRFAMs; TIGR01179; gale; 1.
SQ SEQUENCE 328 AA; 36791 MW; FA49B7EC02793AA CRC64;

Query Match          50.08; Score 48; DB 2; Length 328;
Best Local Similarity 55.68; Pred. No. 7.5;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Oy 1 KKVIAAPRLPPDQILKSD 18
    | : |||| | : ||
Db 280 KVELAPRRAGDPSVLISD 297

RESULT 3
ID O9PNG3 PRELIMINARY; PRT; 328 AA.
AC O9PNG3;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE UDP-glucose 4-epimerase (EC 5.1.3.2).
GN GALE OR C11131C.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCITC 11168;
RX MEDLINE=20150912; PubMed=10688204;
```

```
RA Parkhill J., Wren B.W., Mungall K., Kettle J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
RA Jagers K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrell B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences."
RL Nature 403:665-668(2000).
DR EMBL; AL139077; CAB73386.1; -.
DR HSSP; P09147; 1KVS.
DR InterPro; IPR001509; Epimerase_Dh.
DR Pfam; PF01370; Epimerase; 1.
DR TIGRFAMs; TIGR01179; gale; 1.
KW Complete proteome.
SQ SEQUENCE 328 AA; 36734 MW; 7372DDC08F8F5676 CRC64;

Query Match          50.08; Score 48; DB 16; Length 328;
Best Local Similarity 55.68; Pred. No. 7.5;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Oy 1 KKVIAAPRLPPDQILKSD 18
    | : |||| | : ||
Db 280 KVELAPRRAGDPSVLISD 297

RESULT 4
ID O9K5D5 PRELIMINARY; PRT; 330 AA.
AC O9K5D5;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE UDP-glucose 4-epimerase.
GN GALE.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=81116;
RA Fry B.N., Korolik V., van der Zeijst B.A., Coloe P.J.;
RT "A gene cluster from Campylobacter jejuni involved in inner core and
RT lipid A synthesis."
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ131360; CAC01389.1; -.
DR HSSP; P09147; IXEL.
DR InterPro; IPR001509; Epimerase_Dh.
DR Pfam; PF01370; Epimerase; 1.
DR TIGRFAMs; TIGR01179; gale; 1.
SQ SEQUENCE 330 AA; 37053 MW; 143E592891246F9E CRC64;

Query Match          50.08; Score 48; DB 2; Length 330;
Best Local Similarity 55.68; Pred. No. 7.5;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Oy 1 KKVIAAPRLPPDQILKSD 18
    | : |||| | : ||
Db 282 KVELAPRRAGDPSVLISD 299

RESULT 5
ID O9NAK2 PRELIMINARY; PRT; 755 AA.
AC O9NAK2;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 81.7 kDa protein.
GN F54A3.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoides;
OC Rhabditidae; Peloderinae; Caenorhabditis.
```

OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA None;  
 RT "Genome sequence of the nematode *C. elegans*: a platform for  
 RL investigating biology. The *C. elegans* Sequencing Consortium."  
 RL Science 282:2012-2018(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Goela D., Miller N., Latreille P., Scheet P.;  
 RT "The sequence of *C. elegans* cosmid F54A3."  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Waterston R.;  
 RT "Direct Submission."  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC024136; AF35961.1;  
 DR HSSP; P35520; LJBO.  
 DR InterPro; IPR001926; B6\_enzyme\_beta.  
 DR Pfam; PF00291; PALP; 2.  
 KW Hypothetical protein.  
 SQ SEQUENCE 755 AA; 81678 MW; 042908A27D844DD5 CRC64;

Query Match 46.9%; Score 45; DB 5; Length 755;  
 Best Local Similarity 50.0%; Pred. No. 56;  
 Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 KKVIAPRRLPDPQILKSD 18  
 || || | | | | | | | |  
 Db 298 KKPKPREIFDPKVLVDYD 315

## RESULT 6

OX Q8QVC6 PRELIMINARY; PRT; 482 AA.  
 ID Q8QVC6  
 AC Q8QVC6  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE VP7 (Fragment)  
 OS Colorado tick fever virus.  
 OC Viruses; dsRNA viruses; Reoviridae; Coltivirus.  
 OX NCBI\_TaxID=46839;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CTFV-S6-14-03;  
 RX MEDLINE=21934629; PubMed=11958454;  
 RA Attoui H., Mohd J.F., Biagini P., Cantaloube J.F., de M.P.,  
 RT Murphy F.A., de L.X.;  
 RT "Genus Coltivirus (family Reoviridae): genomic and morphologic  
 RL characterization of Old World and New World viruses."  
 Arch. Virol. 147:533-561(2002).  
 DR EMBL; AF343057; AM18363.1;  
 FT NON\_TER 1  
 FT NON\_TER 482  
 SQ SEQUENCE 482 AA; 53907 MW; 354BA204FADA279F CRC64;

Query Match 46.4%; Score 44.5; DB 12; Length 482;  
 Best Local Similarity 52.9%; Pred. No. 43;  
 Matches 9; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

QY 4 IAPRLPDP-QILKSDT 19  
 || | | | | | | | |  
 Db 388 VAPSQAPDPELSDT 404

## RESULT 7

Q9ENK9 PRELIMINARY; PRT; 684 AA.  
 ID Q9ENK9  
 AC Q9ENK9  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE VP7.  
 OS Colorado tick fever virus.  
 OC Viruses; dsRNA viruses; Reoviridae; Coltivirus.  
 OX NCBI\_TaxID=46839;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-FLORIO;  
 RX MEDLINE=20351236; PubMed=10891382;  
 RA Attoui H., Billoir F., Biagini P., Cantaloube J.F., de Chesse R.,  
 RT De Micco P., de Lamballerie X.;  
 RT "Sequence Determination and Analysis of the Full-Length Genome of  
 RL Colorado Tick Fever Virus, the Type Species of Genus Coltivirus  
 (Family Reoviridae)."  
 RL Biochem. Biophys. Res. Commun. 273:1121-1125(2000).  
 DR EMBL; AF139763; AAG00072.1;  
 SQ SEQUENCE 684 AA; 76236 MW; 8FE2516BB344B4C5 CRC64;

Query Match 46.4%; Score 44.5; DB 12; Length 684;  
 Best Local Similarity 52.9%; Pred. No. 61;  
 Matches 9; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

QY 4 IAPRLPDP-QILKSDT 19  
 || | | | | | | | |  
 Db 511 VAPSQAPDPELSDT 527

## RESULT 8

ID Q9H8S1 PRELIMINARY; PRT; 413 AA.  
 AC Q9H8S1  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE OVARC1001154 protein.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-OVARIAN CARCINOMA;  
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sato H., Sugano S., Shiratori A., Sudo H.,  
 RA Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saifu K.,  
 RA Yanamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,  
 RA Ninomiya K., Iwayanagi T.;  
 RT "NEDO human cDNA sequencing project."  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK023348; BAB14535.1;  
 DR InterPro; IPR000118; Granulin.  
 DR InterPro; IPR001211; PhospholipaseA2.  
 DR Pfam; PF00396; granulin; 4.  
 DR SMART; SM00277; GRAN; 4.  
 DR PROSITE; PS00799; GRANULINS; UNKNOWN\_4.  
 DR PROSITE; PS00118; PA2\_HIS; UNKNOWN\_1.  
 SQ SEQUENCE 413 AA; 44132 MW; 0E3767A44BE314EC CRC64;

Query Match 45.8%; Score 44; DB 4; Length 413;  
 Best Local Similarity 55.6%; Pred. No. 44;  
 Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 KKVIAPRRLPDPQILKSD 18  
 || | | | | | | | |  
 Db 166 EKAPHLSPDPQALKRD 183

```

RESULT 9
Q13648 PRELIMINARY; PRT; 463 AA.
ID Q13648
AC Q13648;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Testis specific leucine rich repeat protein.
GN TSLRP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA O'Hern P.A., Yavetz H., Moy T., Yavetz B., Liang Z.G., Wang G.Y.,
RA Goldberg E.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U60666; AAB02976.1; -.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003603; LRRcap.
DR InterPro; IPR003592; LRR_out.
DR Pfam; PF00560; LRR; 3.
DR SMART; SM00370; LRR; 2.
DR SMART; SM00446; LRRcap; 1.
SQ SEQUENCE 463 AA; 53936 MW; 2D6A11C3A7CFC6C3 CRC64;

Query Match 45.8%; Score 44; DB 4; Length 463;
Best Local Similarity 44.4%; Pred. No. 49;
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 KKVIAPRLPDQILKSD 18
Db 431 EKHTRRRPEPKIIPSE 448

RESULT 10
Q9HBT9 PRELIMINARY; PRT; 466 AA.
ID Q9HBT9
AC Q9HBT9;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Testis specific leucine rich repeat protein.
GN TSLRP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Blechschmidt K., Schudy A., Kalaydjieva L., Gooding R., Gresham D.,
RA Baas F., Jonge, Rd, Schillabel M.B., Menzel U., Baumgart C.,
RA Dette M.D., Jahn N., Rosenthal A.;
RT "Chromosome 8 genomic sequence.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Genome-Sequencing Center Jena.
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF216667; AAG33017.1; -.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003603; LRRcap.
DR InterPro; IPR003592; LRR_out.
DR Pfam; PF00560; LRR; 3.
DR SMART; SM00370; LRR; 3.
DR SMART; SM00446; LRRcap; 1.
SQ SEQUENCE 466 AA; 54254 MW; 4A3DBF849D5621C7 CRC64;

Query Match 45.8%; Score 44; DB 4; Length 466;
Best Local Similarity 44.4%; Pred. No. 50;
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 KKVIAPRLPDQILKSD 18
Db 431 EKHTRRRPEPKIIPSE 448

RESULT 11
Q8TNC8 PRELIMINARY; PRT; 1300 AA.
ID Q8TNC8
AC Q8TNC8;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical protein MA2362.
GN MA2362.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Calagan J.E., Nusbäum C., Roy A., Endrizzi M.G., Macdonald P.,
RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., Dearellano K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuettnner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Unayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity.";
RL Genome Res. 12:532-542(2002).
DR EMBL; AE010925; AAM05750.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 1300 AA; 148273 MW; 591A14D4B8AA7EA1 CRC64;

Query Match 45.8%; Score 44; DB 17; Length 1300;
Best Local Similarity 50.0%; Pred. No. 1.5e+02;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KKVIAPRLPDQILKSD 16
Db 535 EKLVAYSRLPPEVK 550

RESULT 12
O64621 PRELIMINARY; PRT; 131 AA.
ID O64621
AC O64621; Q93YP2;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Expressed protein (Hypothetical 14.3 kDa protein).
GN AT2G18910.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=CV. COLUMBIA;
RA Rounsley S.D., Lin X., Ketchum K.A., Crosby M.L., Brandon R.C.,
RA Sykes S.M., Kaul S., Mason T.M., Kerlavage A.R., Adams M.D.,
RA Somerville C.R., Venter J.C.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=CV. COLUMBIA;
SQ SEQUENCE 131 AA; 148273 MW; 591A14D4B8AA7EA1 CRC64;

```

```

RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLOMBIA;
RA Town C.D., Kaul S.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Nguyen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC003673; AAC09038.2; -
DR EMBL: AY059891; AAL24373.1; -
DR EMBL: AY093348; AAM13347.1; -
KW Hypothetical protein.
SQ SEQUENCE 131 AA; 14339 MW; 3A33878DC8169802 CRC64;

Query Match 44.8%; Score 43; DB 10; Length 131;
Best Local Similarity 58.8%; Pred. No. 19;
Matches 10; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 KVIAPRRRLPDPOILKS 17
| | | | | | | | | |
DB 42 KPVMAAPPVLPFPQPFKS 58

RESULT 13
Q99P28 PRELIMINARY; PRT; 164 AA.
AC Q99P28;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE ARG (Fragment).
GN ARG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
GX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Yu H.-H., Zisch A.H., Dodelet V.C., Pasquale E.B.;
RT "Multiple Signaling Interactions of Abl and Arg Kinases with EphB2 Receptor.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF339052; AAK08120.1; -
FT NON_TER 164 164
SQ SEQUENCE 164 AA; 17029 MW; DE8B75014ACF272 CRC64;

Query Match 44.8%; Score 43; DB 11; Length 164;
Best Local Similarity 57.1%; Pred. No. 24;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 PRELPDPQILKSDT 19
| | | | | | | |
DB 15 PLRAPDPATESDS 28

```

```

RESULT 14
Q40159 PRELIMINARY; PRT; 174 AA.
ID Q40159;
AC Q40159;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Late embryogenesis (Lea)-like protein.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
GX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MARMANDE; TISSUE-MELOIDOGYNE INCOGNITA-INDUCED ROOT KNOT;
RX MEDLINE-96158345; PubMed-8580972;
RA Van der Eycken W., De Almeida Engler J., Inze D., Van Montagu M.,
RA Cheysen G.;
RT "A molecular study of Meloidogyne incognita-induced feeding sites.";
RL Plant J. 9:45-54(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-MARMANDE; TISSUE-MELOIDOGYNE INCOGNITA-INDUCED ROOT KNOT;
RA Van der Eycken W.V.; to the EMBL/GenBank/DBJ databases.
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z46654; CAA86613.1; -
DR InterPro; IPR004864; LEA_2.
DR Pfam; PF03168; LEA_2; 1.
SQ SEQUENCE 174 AA; 19123 MW; 1DBB432943707322 CRC64;

Query Match 44.8%; Score 43; DB 10; Length 174;
Best Local Similarity 47.1%; Pred. No. 26;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 2 KVIAPRRRLPDPOILKSD 18
| | | | | | | | | |
DB 68 RVIASGRIPDPGSIRAN 84

RESULT 15
O80056 PRELIMINARY; PRT; 223 AA.
ID O80056;
AC O80056;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Orf 17.
OS Staphylococcus aureus bacteriophage PVL.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
GX NCBI_TaxID=71366;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-98067870; PubMed-9404084;
RA Kaneko J., Muramoto K., Kamio Y.;
RT "Gene of LukP-PV-like component of Pantone-Valentine leukocidin in Staphylococcus aureus P83 is linked with lukM.";
RL Biosci. Biotechnol. Biochem. 61:1960-1962(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-98332719; PubMed-9666077;
RA Kaneko J., Kimura T., Narita S., Tomita T., Kamio Y.;
RT "Complete nucleotide sequence and molecular characterization of the temperate staphylococcal bacteriophage phi PVL carrying Pantone-Valentine leukocidin genes.";
RL Gene 215:57-67(1998).
DR EMBL: AB009866; BAA31890.1; -
SQ SEQUENCE 223 AA; 25197 MW; 3F91EE45EA9C5638 CRC64;

Query Match 44.8%; Score 43; DB 9; Length 223;
Best Local Similarity 57.1%; Pred. No. 34;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

```

QY 6 PRRLPDPQILKSDT 19  
| | | | |  
Db 141 PNNLEPNSISNSDT 154

Search completed: July 8, 2003, 16:27:56  
Job time.: 43.9231 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2003, 16:08:55 ; Search time 30 Seconds  
(without alignments)

53,300 Million cell updates/sec

Title: US-09-824-647-4

Perfect score: 69

Sequence: 1 PDAKTQCPDST 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	69	100.0	12	AAW85477	Mouse GP88 autocri
2	69	100.0	589	AA14325	Rat epithelin prec
3	69	100.0	589	AA14327	Mouse epithelin pr
4	69	100.0	589	AAW85474	Mouse GP88 autocri
5	69	100.0	589	AAE20520	Mouse granulin/epi
6	51	73.9	56	AA141880	Granulin B. Homo
7	51	73.9	593	AA14326	Human epithelin pr
8	51	73.9	593	AA148673	Granulin sequence.
9	51	73.9	593	AAW85475	Human GP88 autocri
10	51	73.9	593	AAE20521	Human granulin/epi

11	51	73.9	621	21	AB433971	Human cancer assoc
12	43	62.3	57	14	AA41886	Granulin F. Homo
13	41	59.4	565	18	AAW01792	Rat CNKI protein k
14	40	58.0	121	22	AAU61612	Propionibacterium
15	39	56.5	323	22	AAU42060	Propionibacterium
16	38	55.1	37	22	AB339824	Peptide #7330 enco
17	38	55.1	37	22	ABM60551	Human brain expres
18	38	55.1	37	22	AAW73207	Human bone marrow
19	38	55.1	37	22	AAW33416	Peptide #7453 enco
20	38	55.1	37	23	ABG43054	Human peptide enco
21	38	55.1	94	22	AB987388	Human gene 47 enco
22	38	55.1	94	23	ABG65366	Human albumin fusi
23	38	55.1	234	22	AB202888	Vesicular stomatit
24	38	55.1	258	23	ABP28856	Streptococcus poly
25	38	55.1	322	22	ABG15111	Novel human diagno
26	38	55.1	339	17	AAW97754	Canine IgA. Canis
27	38	55.1	351	23	AB449837	Listeria monocytog
28	38	55.1	446	22	ABG03476	Novel human diagno
29	38	55.1	511	22	AB202920	Vesicular stomatit
30	37	53.6	118	21	AA16327	Eucalyptus grandis
31	37	53.6	179	12	AA15426	Bovine epithelin p
32	37	53.6	318	21	AAV93769	Amino acid sequenc
33	37	53.6	398	22	AB31974	Peptide #4625 enco
34	37	53.6	398	22	AB37219	Peptide #4725 enco
35	37	53.6	398	22	AB22520	Protein #4519 enco
36	37	53.6	398	22	AAW57928	Human brain expres
37	37	53.6	398	22	AAW70351	Human bone marrow
38	37	53.6	398	22	AA18190	Peptide #4624 enco
39	37	53.6	398	22	AAW30687	Peptide #4724 enco
40	37	53.6	398	22	AAW05811	Peptide #4493 enco
41	37	53.6	398	23	ABG39996	Human peptide enco
42	37	53.6	474	21	AAW31665	Arabidopsis thalia
43	37	53.6	511	22	AB20289	Vesicular stomatit
44	37	53.6	541	21	AAW31664	Arabidopsis thalia
45	37	53.6	583	21	AAW31663	Arabidopsis thalia

#### ALIGNMENTS

#### RESULT 1

AAW85477  
ID AAW85477 standard; Peptide; 12 AA.

AC AAW85477;

XX 15-MAR-1999 (first entry)

DE Mouse GP88 autocrine growth factor antigenic peptide p12T.

XX GP88; granulin; epithelin; mouse; growth factor; autocrine; tumour;  
KW cancer; viral infection; antagonist; therapy; diagnosis; antigen;  
KW antibody.

XX Mus sp.

PN WO9852607-A1.

XX 26-NOV-1998.

PD 22-MAY-1998; 98WO-US10555.

PF 16-DEC-1997; 97US-0991862.

PR 23-MAY-1997; 97US-0863079.

XX (SERR/) SERRERO G.

PI Serrero G;

XX WPI; 1999-045276/04.

DR Composition containing antagonist of growth factor GP88 - useful for  
XX treating cancer and viral diseases and also for diagnosing disease  
PT



```

PF 03-APR-1991; 91WO-US02321.
XX
PR 13-MAR-1991; 91US-0083796.
PR 03-APR-1990; 90US-0504508.
XX
PA (BRIM ) BRISTOL-MYERS SQUIB.
PI Shoyab M, Plowman GD;
XX WPI: 1991-325168/44.
DR N-PSDB; AAQ14338.
XX
DR New cysteine-rich growth modulating proteins, epithelins - useful
PT as inhibitors of neoplastic cell growth and to promote wound
PT healing and treat psoriasis
XX
PS Disclosure; Fig 18; 97pp; English.
XX
CC ET-1 is a bifunctional growth regulator, capable of stimulating
CC the growth of some cell types while inhibiting the growth of others.
CC ET-2 is functionally similar to ET-1 w.r.t. growth inhibitory
CC bioactivity. In contrast, however, ET-2 is apparently not capable of
CC eliciting the growth stimulatory activity characteristic of ET-1 and,
CC in fact, antagonises this ET-1 activity.
CC See also AAQ14338-40, AAQ14952-53, AAR14328-9 and AAR15315-20.
XX
SQ Sequence 589 AA;

Query Match 100.0%; Score 69; DB 20; Length 12;
Best Local Similarity 100.0%; Pred. No. 7.9e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDAKTCPPDST 12
DB |||||

RESULT 2
AAR14325
ID AAR14325 standard; Protein; 589 AA.
XX
AC AAR14325;
XX
DT 17-JAN-1992 (first entry)
XX
DE Rat epithelin precursor.
XX
KW ET; growth regulation; inhibition; stimulation.
XX
OS Rattus rattus.
XX
FH Key Location/Qualifiers
FT Protein 1..589
FT /label= precursor
FT /note= "claim 11, page 54"
FT Protein 280..335
FT /label= EP-1
FT /note= "claim 12, page 54"
FT Protein 205..261
FT /label= EP-2
FT /note= "claim 13, page 54"
FT Peptide 59..114
FT /label= EP
FT /note= "claim 14, page 54"
FT Peptide 123..179
FT /label= EP
FT /note= "claim 15, page 54"
FT Peptide 362..416
FT /label= EP
FT /note= "claim 16, page 54"
FT Peptide 440..495
FT /label= EP
FT /note= "claim 17, page 54"
FT Peptide 515..570
FT /label= EP
FT /note= "claim 18, page 55"
XX
XX WO9115510-A.
XX
XX 17-OCT-1991.
XX

```

FT Peptide 515..570  
 FT /label= EP  
 FT /note= "claim 28, page 56"  
 XX  
 PN WO9115510-A.  
 XX  
 PD 17-OCT-1991.  
 XX  
 XX  
 XX 03-APR-1991; 91WO-US02321.  
 XX  
 XX 13-MAR-1991; 91US-0083796.  
 PR 03-APR-1990; 90US-0504508.  
 XX  
 XX (BRIM ) BRISTOL-MYERS SQUIB.  
 XX  
 XX Shoyab M, Plowman GD;  
 XX  
 XX WPI; 1991-325168/44.  
 DR N-PSDB; AAQ14340.  
 XX  
 XX New cysteine-rich growth modulating proteins, epithelins - useful  
 PT as inhibitors of neoplastic cell growth and to promote wound  
 PT healing and treat psoriasis  
 PT  
 XX Disclosure; Fig 23; 97pp; English.  
 PS  
 XX ET-1 is a bifunctional growth regulator, capable of stimulating  
 CC the growth of some cell types while inhibiting the growth of others.  
 CC ET-2 is functionally similar to ET-1 w.r.t. growth inhibitory  
 CC bioactivity. In contrast, however, ET-2 is apparently not capable of  
 CC eliciting the growth stimulatory activity characteristic of ET-1 and,  
 CC in fact, antagonises this ET-1 activity.  
 CC See also AAQ14338-40, AAQ14952-53, AAR14328-9 and AAR15315-20.  
 XX  
 XX Sequence 589 AA;  
 SQ  
 Query Match 100.0%; Score 69; DB 12; Length 589;  
 Best Local Similarity 100.0%; Pred. No. 0.0048;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 PDAKTCPPDST 12  
 DB 208 PDAKTCPPDST 219  
 RESULT 4  
 AAQ85474  
 ID AAW85474 standard; Protein; 589 AA.  
 XX  
 XX AAW85474;  
 XX  
 XX 15-MAR-1999 (first entry)  
 DT  
 XX Mouse GP88 autocrine growth factor.  
 DE  
 XX GP88; granulin; epithelin; mouse; growth factor; autocrine; tumour;  
 KW cancer; viral infection; antagonist; therapy; diagnosis.  
 KW  
 XX Mus sp.  
 OS  
 XX Key Location/Qualifiers  
 FH Misc-difference 8 /note= "encoded by CTG"  
 FT  
 FT Misc-difference 54 /note= "encoded by AGC"  
 FT  
 FT Misc-difference 377 /note= "encoded by TGA"  
 FT  
 FT Peptide 208..219  
 FT /note= "p12T peptide used to raise antibody"  
 FT  
 FT Peptide 344..362  
 FT /note= "K19T peptide, used to raise antibody"  
 FT  
 FT Peptide 562..575  
 FT /note= "S14R peptide, used to raise antibody"

XX  
 PN WO9852607-A1.  
 XX  
 XX 26-NOV-1998.  
 XX  
 XX 22-MAY-1998; 98WO-US10555.  
 PF  
 XX 16-DEC-1997; 97US-0991862.  
 PR 23-MAY-1997; 97US-0863079.  
 XX  
 XX (SERR/) SERRERO G.  
 PA  
 XX Serrero G;  
 PI  
 XX WPI; 1999-045276/04.  
 DR N-PSDB; AAV82824.  
 DR  
 XX Composition containing antagonist of growth factor GP88 - useful for  
 PT treating cancer and viral diseases and also for diagnosing disease  
 PT from altered gp88 expression  
 PT  
 XX Example 5; Fig 3A-D; 86pp; English.  
 PS  
 XX This is the amino acid sequence of murine GP88, an 88 kDa  
 CC glycoprotein autocrine growth factor and epithelin/granulin  
 CC precursor that is expressed in a tightly regulated manner in normal  
 CC cells, is overexpressed and unregulated in highly tumorigenic cells  
 CC derived from normal cells, and which acts as a stringently required  
 CC growth stimulator for the tumorigenic cells. Inhibition of GP88  
 CC expression or action in the tumorigenic cells results in an  
 CC inhibition of the tumorigenic properties of the overproducing  
 CC cells. Murine GP88 cDNA (see AAV82824) was obtained from the highly  
 CC tumorigenic PC cell line. Antagonists to GP88 are used to treat  
 CC diseases associated with increased expression of GP88, particularly  
 CC cancer but also viral infections. Fragments of GP88 are used to  
 CC raise specific antibodies (used as antagonists, as diagnostic  
 CC reagents and for delivering toxins or other compounds to GP88-  
 CC expressing cells) and to screen for antibodies. Methods are  
 CC provided for diagnosing disease, or determining susceptibility to  
 CC disease, resulting from altered GP88 activity.  
 XX  
 XX Sequence 589 AA;  
 SQ  
 Query Match 100.0%; Score 69; DB 20; Length 589;  
 Best Local Similarity 100.0%; Pred. No. 0.0048;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 PDAKTCPPDST 12  
 DB 208 PDAKTCPPDST 219  
 RESULT 5  
 AAE20520  
 ID AAE20520 standard; Protein; 589 AA.  
 XX  
 XX AAE20520;  
 XX  
 XX 01-JUL-2002 (first entry)  
 DT  
 XX Mouse granulin/epithelin precursor (GP88) protein.  
 DE  
 XX Mouse; granulin precursor; GP88; cytostatic; tumourigenicity; tamoxifen;  
 KW antineoplastic; antioestrogen therapy; skin cancer.  
 KW  
 XX Mus sp.  
 OS  
 XX Key Location/Qualifiers  
 FH Misc-difference 335 /label= Unknown  
 FT  
 FT /note= "Encoded by ATG"  
 FT 344..362  
 FT  
 FT /note= "Regions used as antigens to raise anti-GP88"

FT Region antibodies  
 FT 562..575  
 FT /note= "Regions used as antigens to raise anti-GP88  
 FT antibodies"  
 FT Misc-difference 586  
 FT /note= "Encoded by ACA"  
 XX  
 PN US2002025543-A1.  
 XX  
 PD 28-FEB-2002.  
 XX  
 PF 15-JUN-2001; 2001US-0880842.  
 XX  
 PR 23-MAY-1997; 97US-0863079.  
 PR 08-DEC-1999; 99US-0456886.  
 XX  
 PA (SERR/) SERRERO G.  
 XX  
 PI Serrero G;  
 XX  
 DR WPI: 2002-267529/31.  
 DR N-PSDB; AAD32849.  
 XX  
 PT Diagnosing tumorigenicity in a human, comprising obtaining a cell  
 PT sample, detecting GP88 in the cells, and determining the number of GP88  
 PT positive cells in the sample.  
 XX  
 PS Disclosure; Fig 8; 50pp; English.  
 XX  
 CC The invention relates to a method for diagnosing tumorigenicity in a  
 CC human. The method comprises obtaining a biological sample containing  
 CC cells from the patient, detecting GP88 in the cells of the sample,  
 CC and determining the number of GP88 positive cells in the sample, and  
 CC determining the ratio of GP88 positive cells to the total number of cells  
 CC in the sample. The invention also relates to a method for determining  
 CC if a human patient is resistant to the antineoplastic effects of  
 CC antioestrogen therapy. The method is useful for diagnosing  
 CC tumorigenicity in a sample, such as blood, serum, plasma, urine, nipple  
 CC aspirate, cerebrospinal fluid, liver, kidney, breast, bone, brain, lung,  
 CC colon, or skin cancer. The method can be used to treat or prevent  
 CC re-occurrence of cancer in a patient, by administering tamoxifen if the  
 CC sample contains less than 10 % GP88, or less than 5 % GP88 positive  
 CC cells. The present sequence is mouse granulins/epithelin precursor (GP89)  
 CC protein.  
 XX  
 XX Sequence 589 AA;  
 SQ  
 Query Match 100.0%; Score 69; DB 23; Length 589;  
 Best Local Similarity 100.0%; Pred. NO. 0.0048;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 PDAKTCPPDST 12  
 Db 208 PDAKTCPPDST 219  
 RESULT 6  
 AAR41880  
 ID - AAR41880 standard; peptide; 56 AA.  
 XX  
 AC AAR41880;  
 XX  
 DT 22-APR-1994 (first entry)  
 DE Granulin B.  
 XX  
 KW Granulin; keratinocytes; wound healing; inhibition; peptide;  
 KW granulocytes; leucocytes.  
 OS Homo sapiens.  
 XX  
 PN WO9315195-A.  
 XX

PD 05-AUG-1993.  
 XX  
 PF 28-FEB-1992; 92WO-CA00089.  
 XX  
 PR 03-FEB-1992; 92US-0829233.  
 XX  
 PA (SOLO/) SOLOMON S.  
 XX  
 PI Solomon S;  
 XX  
 DR WPI: 1993-320328/40.  
 XX  
 PF New cystine rich granulin peptide(s) from leucocyte(s) - are  
 PF keratinocyte inhibitors useful topically for wound healing  
 XX  
 PS Claim 3; Page 28; 53pp; English.  
 XX  
 CC The granulin inhibits keratinocytes and is useful in formulations  
 CC for promoting the healing of wounds. This peptide was designated  
 CC granulin B.  
 XX  
 SQ Sequence 56 AA;  
 Query Match 73.9%; Score 51; DB 14; Length 56;  
 Best Local Similarity 66.7%; Pred. NO. 0.34;  
 Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 PDAKTCPPDST 12  
 Db 4 PDARSRCPDST 15  
 RESULT 7  
 AAR14326  
 ID - AAR14326 standard; Protein; 593 AA.  
 XX  
 AC AAR14326;  
 XX  
 DT 17-JAN-1992 (first entry)  
 DE Human epithelin precursor.  
 XX  
 KW ET; growth regulation; inhibition; stimulation.  
 OS Homo sapiens.  
 XX  
 FH Key  
 FT Location/Qualifiers  
 FT 1..593  
 FT /label= precursor  
 FT /note= "claim 1, page 53"  
 FT 282..337  
 FT /label= EP-1  
 FT /note= "claim 2, page 53"  
 FT 206..262  
 FT /label= EP-2  
 FT /note= "claim 3, page 53"  
 FT 59..114  
 FT /label= EP  
 FT /note= "claim 4, page 53"  
 FT 124..180  
 FT /label= EP  
 FT /note= "claim 5, page 53"  
 FT 364..418  
 FT /label= EP  
 FT /note= "claim 6, page 53"  
 FT 442..497  
 FT /label= EP  
 FT /note= "claim 7, page 53"  
 FT 519..574  
 FT /label= EP  
 FT /note= "claim 8, page 55"  
 PN WO9115510-A.

XX 17-OCT-1991.  
PD 03-APR-1991; 91WO-US02321.  
XX 13-MAR-1991; 91US-0083796.  
PR 03-APR-1990; 90US-0504508.  
XX (BRIM ) BRISTOL-MYERS SQUIB.  
PA Shoyab M, Plowman GD;  
PI WPI; 1991-325168/44.  
XX N-PSDB; AAQ14339.  
DR New cysteine-rich growth modulating proteins, epithelins - useful  
PT as inhibitors of neoplastic cell growth and to promote wound  
PT healing and treat psoriasis  
XX Disclosure; Fig 22; 97pp; English.  
PS ET-1 is a bifunctional growth regulator, capable of stimulating  
XX the growth of some cell types while inhibiting the growth of others.  
CC ET-2 is functionally similar to ET-1 w.r.t. growth inhibitory  
CC bioactivity. In contrast, however, ET-2 is apparently not capable of  
CC eliciting the growth stimulatory activity characteristic of ET-1 and,  
CC in fact, antagonises this ET-1 activity.  
CC See also AAQ14338-40, AAQ14952-53, AAR14328-9 and AAR15315-20.  
XX Sequence 593 AA;  
XX Query Match 73.9%; Score 51; DB 12; Length 593;  
Best Local Similarity 66.7%; Pred. No. 4.1;  
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 1 PDAKTQCPDDST 12  
DB 209 PDARSRCPDGST 220  
RESULT 8  
AAR48673  
ID AAR48673 standard; Protein; 593 AA.  
XX AC AAR48673;  
XX 22-APR-1994 (first entry)  
XX Granulin sequence.  
XX Granulin; keratinocytes; wound healing; inhibition; peptide;  
KW granulocytes; leucocytes.  
XX Homo sapiens.  
XX Key Location/Qualifiers  
FH Misc-difference 452 /note= "valine encoded by ATG."  
FT Misc-difference 539 /note= "glycine encoded by CAG."  
FT WO9315195-A.  
XX .05-AUG-1993.  
XX 28-FEB-1992; 92WO-CA00089.  
XX 03-FEB-1992; 92US-0829233.  
XX (SOLO/) SOLOMON S.  
XX Solomon S;  
PI

DR WPI; 1993-320328/40.  
DR N-PSDB; AAQ49052.  
XX New cysteine rich granulin peptide(s) from leucocyte(s) - are  
PT keratinocyte inhibitors useful topically for wound healing  
XX Disclosure; Figure 4c; 53pp; English.  
XX The granulin inhibits keratinocytes and is useful in formulations  
CC for promoting the healing of wounds.  
XX Sequence 593 AA;  
XX Query Match 73.9%; Score 51; DB 14; Length 593;  
Best Local Similarity 66.7%; Pred. No. 4.1;  
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 1 PDAKTQCPDDST 12  
DB 209 PDARSRCPDGST 220  
RESULT 9  
AAW85475  
ID AAW85475 standard; Protein; 593 AA.  
XX AC AAW85475;  
XX 15-MAR-1999 (first entry)  
XX Human GP88 autocrine growth factor.  
XX GP88; granulin; epithelin; human; growth factor; autocrine; tumour;  
KW cancer; viral infection; antagonist; therapy; diagnosis.  
XX Homo sapiens.  
XX Key Location/Qualifiers  
FH Peptide 340..364  
FT /note= "E19V peptide used to raise antibody"  
FT Peptide 566..579  
FT /note= "A14R peptide used to raise antibody"  
XX WO9852607-A1.  
XX 26-NOV-1998.  
XX 22-MAY-1998; 98WO-US10555.  
XX 16-DEC-1997; 97US-0991862.  
PR 23-MAY-1997; 97US-0863079.  
XX (SERR/) SERRERO G.  
XX Serrero G;  
XX WPI; 1999-045276/04.  
DR N-PSDB; AAW82825.  
XX Composition containing antagonist of growth factor GP88 - useful for  
PT treating cancer and viral diseases and also for diagnosing disease  
PT from altered GP88 expression  
XX Example 5; Fig 9A; 86pp; English.  
XX This is the amino acid sequence of human GP88, an 88 kDa  
CC glycoprotein autocrine growth factor and epithelin/granulin  
CC precursor that is expressed in a tightly regulated manner in normal  
CC cells, is overexpressed and unregulated in highly tumorigenic cells  
CC derived from normal cells, and which acts as a stringently required  
CC growth stimulator for the tumorigenic cells. Inhibition of GP88  
CC expression or action in the tumorigenic cells results in an  
CC inhibition of the tumorigenic properties of the overproducing

CC cells. Antagonists to GP88 are used to treat diseases associated  
 CC with increased expression of GP88, particularly cancer but also  
 CC viral infections. Fragments of GP88 are used to raise specific  
 CC antibodies (used as antagonists, as diagnostic reagents and for  
 CC delivering toxins or other compounds to GP88-expressing cells) and  
 CC to screen for antibodies. Methods are provided for diagnosing  
 CC disease, or determining susceptibility to disease, resulting from  
 CC altered GP88 activity.

XX  
 XX  
 SQ Sequence 593 AA;  
 Query Match 73.9%; Score 51; DB 20; Length 593;  
 Best Local Similarity 66.7%; Pred. NO. 4.1;  
 Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 PDAKTCPCDDST 12  
 |||:::|||||  
 Db 209 PDARSRCPDGST 220

RESULT 10  
 AAE20521  
 ID AAE20521 standard; Protein; 593 AA.  
 XX  
 AC AAE20521;  
 XX  
 DT 01-JUL-2002 (first entry)  
 XX  
 DE Human granulin/epithelin precursor (GP88) protein.  
 XX  
 KW Human; granulin precursor; GP88; cytostatic; tumourigenicity; tamoxifen;  
 KW antineoplastic; antioestrogen therapy; skin cancer.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 273  
 FT /note= "Encoded by AAG"  
 FT Region 346..364  
 FT /note= "Region used as antigen to develop anti-human  
 FT GP88 neutralising antibody"

XX US200202543-A1.  
 XX  
 XX 28-FEB-2002.  
 XX  
 XX 15-JUN-2001; 2001US-0880842.  
 XX  
 XX 23-MAY-1997; 97US-0863079.  
 XX 08-DEC-1999; 99US-0456886.  
 XX (SERR/) SERRERO G.  
 XX  
 XX Serrero G;  
 PI  
 XX WPI; 2002-267529/31.  
 DR N-PSDB; AAD32850.  
 XX  
 PT Diagnosing tumorigenicity in a human, comprising obtaining a cell  
 PT sample, detecting GP88 in the cells, and determining the number of GP88  
 PT positive cells in the sample -  
 XX  
 PS Disclosure; Fig 9B; 50pp; English.

XX The invention relates to a method for diagnosing tumorigenicity in a  
 CC human. The method comprises obtaining a biological sample containing  
 CC cells from the patient, detecting GP88 in the cells of the sample,  
 CC and determining the number of GP88 positive cells in the sample, and  
 CC determining the ratio of GP88 positive cells to the total number of cells  
 CC in the sample. The invention also relates to a method for determining  
 CC if a human patient is resistant to the antineoplastic effects of  
 CC antioestrogen therapy. The method is useful for diagnosing  
 CC tumorigenicity in a sample, such as blood, serum, plasma, urine, nipple

CC aspirate, cerebrospinal fluid, liver, kidney, breast, bone, brain, lung,  
 CC colon, or skin cancer. The method can be used to treat or prevent  
 CC re-occurrence of cancer in a patient, by administering tamoxifen if the  
 CC sample contains less than 10 % GP88, or less than 5 % GP88 positive  
 CC cells. The present sequence is human granulin/epithelin precursor (GP88)  
 CC protein.

XX  
 XX  
 SQ Sequence 593 AA;  
 Query Match 73.9%; Score 51; DB 23; Length 593;  
 Best Local Similarity 66.7%; Pred. NO. 4.1;  
 Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 PDAKTCPCDDST 12  
 |||:::|||||  
 Db 209 PDARSRCPDGST 220

RESULT 11  
 AAB43971  
 ID AAB43971 standard; Protein; 621 AA.

XX  
 AC AAB43971;  
 XX  
 DT 08-FEB-2001 (first entry)  
 XX  
 DE Human cancer associated protein sequence SEQ ID NO:1416.

XX  
 KW Human; cancer associated gene; cancer antigen; detection; cancer;  
 KW diagnosis; cytostatic; proliferative; vulnery; immunomodulator;  
 KW antidiabetic; antiaschmatic; antirheumatic; antithrombotic; antiviral;  
 KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiac;  
 KW dermatological; neuroprotective; thrombolytic; coagulant; neotropic;  
 KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;  
 KW immune disorder; haematopoietic cell disorder; autolimmune disorder;  
 KW allergic reaction; graft versus host disease; organ rejection;  
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;  
 KW neurological disease; drug screening.

XX Homo sapiens.  
 OS  
 XX WO200055350-A1.  
 PN  
 XX  
 XX 21-SEP-2000.  
 XX  
 XX 08-MAR-2000; 2000WO-US05882.  
 XX  
 XX 12-MAR-1999; 99US-0124270.  
 PR  
 XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;  
 PI  
 XX WPI; 2000-587533/55.  
 DR N-PSDB; AAC78180.  
 XX

XX Novel isolated nucleic acids comprising sequences encoding peptides  
 XX useful for treating or diagnosing e.g. cancer -

PS Claim 11; Page 2094-2096; 2352pp; English.

XX AAC77607 to AAC78448 encode the human cancer associated proteins given  
 CC in AAB43998 to AAB44239. The proteins can have activities based on the  
 CC tissues and cells the genes are expressed in. Example of activities  
 CC include: cytostatic; proliferative; vulnery; immunomodulator;  
 CC antidiabetic; antiaschmatic; antirheumatic; antithrombotic; antiviral;  
 CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;  
 CC dermatological; neuroprotective; cardiac; thrombolytic; coagulant;  
 CC neotropic; vasotropic; antipsoriatic and antiangiogenic. The  
 CC polynucleotides and polypeptides can be used for preventing, treating or  
 CC ameliorating medical conditions and diagnosing pathological conditions.  
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from  
 CC the present invention may be used to treat immune disorders by activating

or inhibiting the proliferation, differentiation or mobilisation of immune cells, to treat disorders of haematopoietic cells, autoimmune disorders, allergic reactions, graft versus host disease and organ rejection, modulate haemostatic or thrombolytic activity, modulate inflammation, cancers, cardiovascular disorders, neurological disease and bacterial or viral infections. The peptides, nucleotides, antibodies, agonists and antagonists may be also be used in drug screens. AAC78449 to AAC78457 and AAB44240 represent sequences used in the exemplification of the present invention.

SQ Sequence 621 AA;

Query Match 73.9%; Score 51; DB 21; Length 621;

Best Local Similarity 66.7%; Pred. NO. 4.3;

Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 PDAKTQCPDDST 12

||||:|||||

Db 237 PDARSRCPDGST 248

RESULT 12

AAR41886

ID AAR41886 standard; peptide; 57 AA.

XX AC AAR41886;

DT 22-APR-1994 (first entry)

DE Granulin F.

XX Granulin; keratinocytes; wound healing; inhibition; peptide;

KW Granulocytes; leucocytes.

XX Homo sapiens.

OS WO9315195-A.

PN 05-AUG-1993.

PD 28-FEB-1992; 92WO-CA00089.

PF 03-FEB-1992; 92US-0829233.

PR (SOLO/) SOLOMON S.

XX Solomon S;

XX WPI; 1993-320328/40.

XX New cystine rich granulin peptide(s) from leucocyte(s) - are

PT keratinocyte inhibitors useful topically for wound healing

XX Claim 9; Page 34; 53pp; English.

PS The granulin inhibits keratinocytes and is useful in formulations

CC for promoting the healing of wounds. This peptide was designated

CC granulin F.

XX Sequence 57 AA;

Query Match 62.3%; Score 43; DB 14; Length 57;

Best Local Similarity 58.3%; Pred. NO. 7;

Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 PDAKTQCPDDST 12

||||:|||||

Db 5 PDSQFECDFST 16

RESULT 13

AAW01792

ID AAW01792 standard; Protein; 565 AA.

XX AAW01792;

DT 30-SEP-1997 (first entry)

DE Rat CNK1 protein kinase.

XX Protein kinase; enzymatic; Saccharomyces cerevisiae; CDC7;

KW cell cycle regulatory protein.

XX Rattus rattus.

OS JP08322571-A.

PN 10-DEC-1996.

PD 05-JUN-1995; 95JP-0138152.

PF 05-JUN-1995; 95JP-0138152.

PR (SHKJ) SHINGIJUTSU JIGYODAN.

XX WPI; 1997-081085/08.

XX N-PSDB; AAT59400.

DR CNK1 gene encodes polypeptide with protein kinase activity - is

PT homologous to Saccharomyces cerevisiae CDC7 cell cycle regulatory

XX protein

XX Example 1; Page 6; 9pp; Japanese.

PS The present sequence represents the CNK1 protein kinase derived from

XX rat. The gene encoding this sequence is homologous to CDC7 which

CC originates from a mammalian and controls the cell cycle.

XX Sequence 565 AA;

Query Match 59.4%; Score 41; DB 18; Length 565;

Best Local Similarity 50.0%; Pred. NO. 1.7e+02;

Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 PDAKTQCPDDST 12

||||:|||||

Db 375 PEVLTKCPDQFT 386

RESULT 14

AAU61612

ID AAU61612 standard; Protein; 121 AA.

XX AAU61612;

XX 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #22508.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;

XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;

KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;

KW dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

OS WO200181581-A2.

PN 01-NOV-2001.

PD 20-APR-2001; 2001WO-US12865.

PF 21-APR-2000; 2000US-199047P.

XX 02-JUN-2000; 2000US-208841P.

PR 07-JUL-2000; 2000US-216747P.

XX

(CORI-) CORIXA CORP.  
 PA Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;  
 XX  
 DR WPI; 2001-616774/71.  
 DR N-PSDB; AAS59619.  
 XX  
 PT Propionibacterium acnes polypeptides and nucleic acids useful for  
 PT vaccinating against and diagnosing infections, especially useful for  
 PT treating acne vulgaris -  
 XX  
 PS Example 1; SEQ ID NO 22807; 1069pp; English.  
 XX  
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
 CC polypeptides. The proteins and their associated DNA sequences are used in  
 CC the treatment, prevention and diagnosis of medical conditions caused by  
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.  
 CC P. acnes is also involved in infections of bone, joints and the central  
 CC nervous system, however it is particularly involved in the inflammatory  
 CC lesions associated with acne vulgaris. A method for detecting the  
 CC presence or absence of P. acnes in a patient comprises contacting a  
 CC sample with a binding agent that binds to the proteins of the invention  
 CC and determining the amount of bound protein in the sample. The  
 CC polypeptides may be used as antigens in the production of antibodies  
 CC specific for P. acnes proteins. These antibodies can be used to  
 CC downregulate expression and activity of P. acnes polypeptides and  
 CC therefore treat P. acnes infections. The antibodies may also be used as  
 CC diagnostic agents for determining P. acnes presence, for example, by  
 CC enzyme linked immunosorbent assay (ELISA).  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 121 AA;  
 Query Match 58.0%; Score 40; DB 22; Length 121;  
 Best Local Similarity 60.0%; Pred. No. 47;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 PDAKTCQCPDD 10  
 | | : |||||  
 DB 68 FCARRRCPPD 77  
 RESULT 15  
 AAU42060  
 ID AAU42060 standard; Protein; 323 AA.  
 XX  
 AC AAU42060;  
 XX  
 DT 27-FEB-2002 (first entry)  
 XX  
 DE Propionibacterium acnes immunogenic protein #2956.  
 XX  
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
 KW dermatological; osteopathic; neuroprotectant.  
 XX  
 OS Propionibacterium acnes.  
 XX  
 PN WO200181581-A2.  
 XX  
 PD 01-NOV-2001.  
 XX  
 PF 20-APR-2001; 2001WO-US12865.  
 XX  
 PR 21-APR-2000; 2000US-199047P.  
 PR 02-JUN-2000; 2000US-208841P.  
 PR 07-JUL-2000; 2000US-216747P.  
 XX

(CORI-) CORIXA CORP.  
 PA Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;  
 XX  
 DR WPI; 2001-616774/71.  
 DR N-PSDB; AAS59516.  
 XX  
 PT Propionibacterium acnes polypeptides and nucleic acids useful for  
 PT vaccinating against and diagnosing infections, especially useful for  
 PT treating acne vulgaris -  
 XX  
 PS Claim 3; SEQ ID NO 3255; 1069pp; English.  
 XX  
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
 CC polypeptides. The proteins and their associated DNA sequences are used in  
 CC the treatment, prevention and diagnosis of medical conditions caused by  
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.  
 CC P. acnes is also involved in infections of bone, joints and the central  
 CC nervous system, however it is particularly involved in the inflammatory  
 CC lesions associated with acne vulgaris. A method for detecting the  
 CC presence or absence of P. acnes in a patient comprises contacting a  
 CC sample with a binding agent that binds to the proteins of the invention  
 CC and determining the amount of bound protein in the sample. The  
 CC polypeptides may be used as antigens in the production of antibodies  
 CC specific for P. acnes proteins. These antibodies can be used to  
 CC downregulate expression and activity of P. acnes polypeptides and  
 CC therefore treat P. acnes infections. The antibodies may also be used as  
 CC diagnostic agents for determining P. acnes presence, for example, by  
 CC enzyme linked immunosorbent assay (ELISA).  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 323 AA;  
 Query Match 56.5%; Score 39; DB 22; Length 323;  
 Best Local Similarity 50.0%; Pred. No. 1.9e+02;  
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 PDAKTCQCPDDST 12  
 | | : |||||  
 DB 39 PEPQLGCPDDSS 50

Search completed: July 8, 2003, 16:24:14  
 Job time : 31 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2003, 16:21:00 ; Search time 7.38462 Seconds  
(without alignments)  
47.812 Million cell updates/sec

Title: US-09-824-647-4  
Perfect score: 69  
Sequence: 1 PDAKTQCPDST 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_AA:\*

- 1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pap:\*
- 2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pap:\*
- 3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pap:\*
- 4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pap:\*
- 5: /cgn2\_6/ptodata/1/1aa/PTDUS\_COMB.pap:\*
- 6: /cgn2\_6/ptodata/1/1aa/backfiles1.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	69	100.0	12	4	US-08-991-862-4
2	69	100.0	589	1	US-07-668-648-2
3	69	100.0	589	1	US-07-668-648-6
4	69	100.0	589	2	US-08-429-998-2
5	69	100.0	589	2	US-08-429-998-6
6	69	100.0	589	2	US-08-431-333-2
7	69	100.0	589	2	US-08-431-333-6
8	69	100.0	589	4	US-08-991-862-2
9	69	100.0	589	5	PT-US91-02321-2
10	69	100.0	589	5	PT-US91-02321-6
11	51	73.9	593	1	US-07-668-648-4
12	51	73.9	593	2	US-08-429-998-4
13	51	73.9	593	2	US-08-431-333-4
14	51	73.9	593	4	US-08-991-862-17
15	51	73.9	593	5	PT-US91-02321-4
16	38	55.1	343	1	US-08-336-891-2
17	38	55.1	343	5	PT-US95-13795-4
18	37	53.6	179	1	US-07-668-648-8
19	37	53.6	179	2	US-08-429-998-8
20	37	53.6	179	5	PT-US91-02321-8
21	37	53.6	179	5	PT-US91-02321-8
22	36	52.2	561	4	US-09-233-989-5
23	36	52.2	1040	4	US-09-564-805-238
24	36	52.2	1480	3	US-09-191-647-7
25	36	52.2	1480	4	US-09-540-245A-7
26	36	52.2	1480	4	US-09-540-153-7
27	36	52.2	1480	4	US-09-182-024A-5

28	36	52.2	1480	5	PCT-0591-09055-2	Sequence 2, Appli
29	35.5	51.4	35	1	US-08-190-802A-220	Sequence 220, App
30	35.5	51.4	35	4	US-08-477-346-220	Sequence 220, App
31	35.5	51.4	35	4	US-08-473-089-220	Sequence 220, App
32	35.5	51.4	35	4	US-08-487-072A-220	Sequence 220, App
33	35.5	51.4	488	1	US-08-190-802A-60	Sequence 60, Appl
34	35.5	51.4	488	4	US-08-477-346-60	Sequence 60, Appl
35	35.5	51.4	488	4	US-08-473-089-60	Sequence 60, Appl
36	35.5	51.4	488	4	US-08-487-072A-60	Sequence 1, Appli
37	35	50.7	388	2	US-08-742-621-1	Sequence 22, Appli
38	35	50.7	388	4	US-09-191-608-22	Sequence 3, Appli
39	35	50.7	407	2	US-08-742-440A-3	Sequence 6, Appli
40	35	50.7	529	3	US-08-821-984-6	Sequence 8, Appli
41	35	50.7	529	3	US-08-821-984-8	Sequence 6, Appli
42	35	50.7	529	4	US-09-329-749-6	Sequence 8, Appli
43	35	50.7	529	4	US-09-329-749-8	Sequence 4, Appli
44	35	50.7	809	1	US-07-789-915A-4	Sequence 4, Appli
45	35	50.7	809	1	US-08-005-002C-4	Sequence 4, Appli

#### ALIGNMENTS

RESULT 1  
US-08-991-862-4  
; Sequence 4, Application US/08991862  
; Patent No. 6309826  
; GENERAL INFORMATION:  
; APPLICANT: Serrero, Ginette  
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
; FILE REFERENCE: Z9996.488/P001-A  
; CURRENT APPLICATION NUMBER: US/08/991,862  
; CURRENT FILING DATE: 1998-08-17  
; EARLIER APPLICATION NUMBER: 08/863,862  
; EARLIER FILING DATE: 1997-05-23  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 12  
; TYPE: PPT  
; ORGANISM: mouse granulin  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(12)  
; OTHER INFORMATION: Internal peptide of mouse GP88 used to raise the  
; OTHER INFORMATION: antisera against the GP88 used in the  
; OTHER INFORMATION: immunoaffinity step.  
US-08-991-862-4

Query Match 100.0%; Score 69; DB 4; Length 12;  
Best Local Similarity 100.0%; Pred. No. 1e-05;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDAKTQCPDST 12  
|||||  
DB 1 PDAKTQCPDST 12

RESULT 2  
US-07-668-648-2  
; Sequence 2, Application US/07668648  
; Patent No. 5416192  
; GENERAL INFORMATION:  
; APPLICANT: Shoyab, Mohammed  
; APPLICANT: Plowman, Gregory D.  
; TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH  
; TITLE OF INVENTION: MODULATING PROTEINS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York



COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/668,648  
FILING DATE: 19910819  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELEPHONE: (212)790-9090  
TELEFAX: (212) 869-9741  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 589 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-668-648-2

Query Match 100.0%; Score 69; DB 1; Length 589;  
Best Local Similarity 100.0%; Pred. No. 0.0007;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDAKTQCPDST 12  
|||||  
DB 208 PDAKTQCPDST 219

RESULT 3  
US-07-668-648-6  
; Sequence 6, Application US/07668648  
; Patent No. 5416192  
; GENERAL INFORMATION:  
; APPLICANT: Shoyab, Mohammed  
; APPLICANT: Plowman, Gregory D.  
; TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH  
; TITLE OF INVENTION: MODULATING PROTEINS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/668,648  
FILING DATE: 19910819  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELEPHONE: (212)790-9090  
TELEFAX: (212) 869-9741  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 589 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear

COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/668,648  
FILING DATE: 19910819  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELEPHONE: (212)790-9090  
TELEFAX: (212) 869-9741  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 589 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-429-998-2

Query Match 100.0%; Score 69; DB 2; Length 589;  
Best Local Similarity 100.0%; Pred. No. 0.0007;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDAKTQCPDST 12  
|||||  
DB 208 PDAKTQCPDST 219

RESULT 4  
US-08-429-998-2  
; Sequence 2, Application US/08429998  
; Patent No. 5885961  
; GENERAL INFORMATION:  
; APPLICANT: Shoyab, Mohammed  
; APPLICANT: Plowman, Gregory D.  
; TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH  
; TITLE OF INVENTION: MODULATING PROTEINS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/429,998  
FILING DATE: 27-APR-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/668,648  
FILING DATE: 13-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELEPHONE: (212)790-9090  
TELEFAX: (212) 869-9741  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 589 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-429-998-2

Query Match 100.0%; Score 69; DB 2; Length 589;  
Best Local Similarity 100.0%; Pred. No. 0.0007;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDAKTQCPDST 12  
|||||  
DB 208 PDAKTQCPDST 219

RESULT 5  
US-08-429-998-6  
; Sequence 6, Application US/08429998  
; Patent No. 5885961  
; GENERAL INFORMATION:  
; APPLICANT: Shoyab, Mohammed  
; APPLICANT: Plowman, Gregory D.  
; TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH

;; TITLE OF INVENTION: MODULATING PROTEINS  
;; NUMBER OF SEQUENCES: 12  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Pennie & Edmonds  
;; STREET: 1155 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/429,998  
;; FILING DATE: 27-APR-1995  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/668,648  
;; FILING DATE: 13-MAR-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Misrock, S. Leslie  
;; REGISTRATION NUMBER: 18,872  
;; REFERENCE/DOCKET NUMBER: 5624-161-999  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212)790-9090  
;; TELEFAX: (212) 869-9741  
;; INFORMATION FOR SEQ ID NO: 6:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 589 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; US-08-429-998-6

Query Match 100.0%; Score 69; DB 2; Length 589;  
Best Local Similarity 100.0%; Pred. No. 0.0007;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDAKTCPPDST 12  
Db 208 PDAKTCPPDST 219

RESULT 6  
US-08-431-333-2  
; Sequence 2, Application US/08431333  
; Patent No. 5965723  
; GENERAL INFORMATION:  
; APPLICANT: Shoyab, Mohammed  
; TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH  
; TITLE OF INVENTION: MODULATING PROTEINS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/431,333  
; FILING DATE: 27-APR-1995  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/668,648

;; FILING DATE: 13-MAR-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Misrock, S. Leslie  
;; REGISTRATION NUMBER: 18,872  
;; REFERENCE/DOCKET NUMBER: 5624-161-999  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212)790-9090  
;; TELEFAX: (212) 869-9741  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 589 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; US-08-431-333-2

Query Match 100.0%; Score 69; DB 2; Length 589;  
Best Local Similarity 100.0%; Pred. No. 0.0007;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDAKTCPPDST 12  
Db 208 PDAKTCPPDST 219

RESULT 7  
US-08-431-333-6  
; Sequence 6, Application US/08431333  
; Patent No. 5965723  
; GENERAL INFORMATION:  
; APPLICANT: Shoyab, Mohammed  
; TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH  
; TITLE OF INVENTION: MODULATING PROTEINS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/431,333  
; FILING DATE: 27-APR-1995  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/668,648  
; FILING DATE: 13-MAR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Misrock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 5624-161-999  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)790-9090  
; TELEFAX: (212) 869-9741  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 589 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-431-333-6

Query Match 100.0%; Score 69; DB 2; Length 589;  
Best Local Similarity 100.0%; Pred. No. 0.0007;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDAKTQCPDDST 12  
Best Local Similarity 100.0%; Score 69; DB 4; Length 589;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 208 PDAKTQCPDDST 219  
RESULT 8  
US-08-991-862-2  
Sequence 2, Application US/08991862  
Patent No. 6309826  
GENERAL INFORMATION:  
APPLICANT: Seriero, Ginette  
TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
FILE REFERENCE: Z9996.488/P001-A  
CURRENT APPLICATION NUMBER: US/08/991,862  
EARLIER FILING DATE: 1998-08-17  
EARLIER APPLICATION NUMBER: 08/863,862  
EARLIER FILING DATE: 1997-05-23  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 2  
LENGTH: 589  
TYPE: PRT  
ORGANISM: Mouse epithelin/granulin  
US-08-991-862-2

Query Match 100.0%; Score 69; DB 4; Length 589;  
Best Local Similarity 100.0%; Pred. No. 0.0007;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDAKTQCPDDST 12  
Best Local Similarity 100.0%; Score 69; DB 5; Length 589;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 208 PDAKTQCPDDST 219  
RESULT 9  
PCT-US91-02321-2  
Sequence 2, Application PC/TUS9102321  
GENERAL INFORMATION:  
APPLICANT: Shoyab, Mohammed  
TITLE OF INVENTION: EPIITHELINS: NOVEL CYSTEINE-RICH GROWTH  
TITLE OF INVENTION: MODULATING PROTEINS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Bristol-Myers Squibb Company  
STREET: 3005 First Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98121  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/02321  
FILING DATE: 19910403  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Poor, Brian W.  
REGISTRATION NUMBER: 32,928  
REFERENCE/DOCKET NUMBER: ON0071A-PC  
TELEPHONE: (206)728-4800  
TELEFAX: (206)448-4775  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 589 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US91-02321-2

Query Match 100.0%; Score 69; DB 5; Length 589;  
Best Local Similarity 100.0%; Pred. No. 0.0007;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDAKTQCPDDST 12  
Best Local Similarity 100.0%; Score 69; DB 5; Length 589;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 208 PDAKTQCPDDST 219  
RESULT 10  
PCT-US91-02321-6  
Sequence 6, Application PC/TUS9102321  
GENERAL INFORMATION:  
APPLICANT: Shoyab, Mohammed  
TITLE OF INVENTION: EPIITHELINS: NOVEL CYSTEINE-RICH GROWTH  
TITLE OF INVENTION: MODULATING PROTEINS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Bristol-Myers Squibb Company  
STREET: 3005 First Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98121  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/02321  
FILING DATE: 19910403  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Poor, Brian W.  
REGISTRATION NUMBER: 32,928  
REFERENCE/DOCKET NUMBER: ON0071A-PC  
TELEPHONE: (206)728-4800  
TELEFAX: (206)448-4775  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 589 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US91-02321-6

Query Match 100.0%; Score 69; DB 5; Length 589;  
Best Local Similarity 100.0%; Pred. No. 0.0007;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDAKTQCPDDST 12  
Best Local Similarity 100.0%; Score 69; DB 5; Length 589;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 208 PDAKTQCPDDST 219  
RESULT 11  
US-07-668-648-4  
Sequence 4, Application US/07668648  
Patent No. 5416192  
GENERAL INFORMATION:  
APPLICANT: Shoyab, Mohammed  
TITLE OF INVENTION: EPIITHELINS: NOVEL CYSTEINE-RICH GROWTH  
TITLE OF INVENTION: MODULATING PROTEINS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York

STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/668,648  
FILING DATE: 19910819  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Misrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELEPHONE: (212)790-9090  
TELEFAX: (212) 869-9741  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 593 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-668-648-4

Query Match 73.9%; Score 51; DB 1; Length 593;  
Best Local Similarity 66.7%; Pred. No. 0.7;  
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 PDAKTCPPDST 12  
Db 209 PDARSRCPDGST 220

RESULT 12  
US-08-429-998-4  
Sequence 4, Application US/08429998  
Patent No. 5885961  
GENERAL INFORMATION:  
APPLICANT: Shoyab, Mohammed  
TITLE OF INVENTION: EPIITHELINS: NOVEL CYSTEINE-RICH GROWTH  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/429,998  
FILING DATE: 27-APR-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/668,648  
FILING DATE: 13-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Misrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELEPHONE: (212)790-9090  
TELEFAX: (212) 869-9741  
INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:  
LENGTH: 593 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-429-998-4

Query Match 73.9%; Score 51; DB 2; Length 593;  
Best Local Similarity 66.7%; Pred. No. 0.7;  
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 PDAKTCPPDST 12  
Db 209 PDARSRCPDGST 220

RESULT 13  
US-08-431-333-4  
Sequence 4, Application US/08431333  
Patent No. 5965723  
GENERAL INFORMATION:  
APPLICANT: Shoyab, Mohammed  
TITLE OF INVENTION: EPIITHELINS: NOVEL CYSTEINE-RICH GROWTH  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/431,333  
FILING DATE: 27-APR-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/668,648  
FILING DATE: 13-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Misrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELEPHONE: (212)790-9090  
TELEFAX: (212) 869-9741  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 593 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-431-333-4

Query Match 73.9%; Score 51; DB 2; Length 593;  
Best Local Similarity 66.7%; Pred. No. 0.7;  
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 PDAKTCPPDST 12  
Db 209 PDARSRCPDGST 220

RESULT 14  
US-08-991-862-17  
Sequence 17, Application US/08991862  
Patent No. 6309826

GENERAL INFORMATION:  
APPLICANT: Seriero, Gnette  
TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
FILE REFERENCE: Z9996.488/P001-A  
CURRENT APPLICATION NUMBER: US/08/991.862  
CURRENT FILING DATE: 1998-08-17  
EARLIER APPLICATION NUMBER: 08/863.862  
EARLIER FILING DATE: 1997-05-23  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 17  
LENGTH: 593  
TYPE: PRT  
ORGANISM: Human GP88 cDNA  
US-08-991-862-17

Query Match 73.9%; Score 51; DB 4; Length 593;

Best Local Similarity 66.7%; Pred. No. 0.7;  
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 PDAKTQCPDDST 12  
Db 209 PDARSRCPDGST 220  
|||||:|||||

## RESULT 15

PCT-US91-02321-4  
Sequence 4, Application PC/TUS9102321  
GENERAL INFORMATION:  
APPLICANT: Shoyab, Mohammed  
APPLICANT: Plowman, Gregory D.  
TITLE OF INVENTION: EPITHELIALS: NOVEL CYSTEINE-RICH GROWTH  
TITLE OF INVENTION: MODULATING PROTEINS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Bristol-Myers Squibb Company  
STREET: 3005 First Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98121  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/02321  
FILING DATE: 19910403  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Poor, Brian W.  
REGISTRATION NUMBER: 32,928  
REFERENCE/DOCKET NUMBER: ON0071A-PC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)728-4800  
TELEFAX: (206)448-4775  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 593 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US91-02321-4

Query Match 73.9%; Score 51; DB 5; Length 593;

Best Local Similarity 66.7%; Pred. No. 0.7;  
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 PDAKTQCPDDST 12  
Db 209 PDARSRCPDGST 220  
|||||:|||||

Search completed: July 8, 2003, 16:30:18  
Job time : 9.38462 secs

Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2003, 16:24:26 ; Search time 11.6923 Seconds  
(without alignments)  
119.483 Million cell updates/sec

Title: US-09-824-647-4  
Perfect score: 69  
Sequence: 1 PDAKTOCPDDST 1

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 116419773 residues

Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pcp.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pcp.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pcp.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pcp.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pcp.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pcp.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pcp.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pcp.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pcp.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pcp.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pcp.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pcp.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pcp.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match					
1	69	100.0	12	9	US-09-824-647-4	Sequence 4, Appli	
2	69	100.0	12	9	US-10-218-509-4	Sequence 4, Appli	
3	69	100.0	12	9	US-10-281-160-4	Sequence 4, Appli	
4	69	100.0	12	10	US-09-813-156-4	Sequence 4, Appli	
5	69	100.0	12	10	US-09-824-807-4	Sequence 2, Appli	
6	69	100.0	589	9	US-09-824-647-2	Sequence 2, Appli	
7	69	100.0	589	9	US-10-218-509-2	Sequence 2, Appli	
8	69	100.0	589	9	US-10-281-160-2	Sequence 2, Appli	
9	69	100.0	589	10	US-09-813-156-2	Sequence 2, Appli	
10	69	100.0	589	10	US-09-824-807-2	Sequence 2, Appli	
11	51	73.9	593	9	US-09-824-647-17	Sequence 17, Appl	
12	51	73.9	593	9	US-10-218-509-17	Sequence 17, Appl	
13	51	73.9	593	9	US-10-281-160-17	Sequence 17, Appl	
14	51	73.9	593	10	US-09-813-156-17	Sequence 17, Appl	
15	51	73.9	593	10	US-09-824-807-17	Sequence 17, Appl	
16	51	73.9	621	10	US-09-825-301-1416	Sequence 1416, Ap	
17	38	55.1	37	10	US-09-864-761-44137	Sequence 44137, A	
18	38	55.1	94	10	US-09-789-561-129	Sequence 129, App	
19	37	53.6	398	10	US-09-864-761-37818	Sequence 37818, A	

20	36	52.2	38	9	US-10-133-128-145	Sequence 145, App
21	36	52.2	132	10	US-09-864-761-37187	Sequence 37187, A
22	36	52.2	197	9	US-09-874-503-4	Sequence 4, Appl
23	36	52.2	197	9	US-10-000-157-4	Sequence 4, Appl
24	36	52.2	197	9	US-10-036-041-11	Sequence 11, Appl
25	36	52.2	197	9	US-09-320-713-29	Sequence 29, Appl
26	36	52.2	197	9	US-09-816-748-4	Sequence 4, Appl
27	36	52.2	197	9	US-09-747-259-4	Sequence 4, Appl
28	36	52.2	197	9	US-10-035-853-11	Sequence 11, Appl
29	36	52.2	197	9	US-10-174-590-448	Sequence 448, App
30	36	52.2	197	9	US-10-176-758-448	Sequence 448, App
31	36	52.2	197	9	US-10-175-737-448	Sequence 448, App
32	36	52.2	197	9	US-10-173-706-448	Sequence 448, App
33	36	52.2	197	9	US-10-175-738-448	Sequence 448, App
34	36	52.2	197	9	US-10-175-752-448	Sequence 448, App
35	36	52.2	197	9	US-10-176-482-448	Sequence 448, App
36	36	52.2	197	9	US-10-176-757-448	Sequence 448, App
37	36	52.2	197	9	US-10-176-913-448	Sequence 448, App
38	36	52.2	197	9	US-10-180-553-448	Sequence 448, App
39	36	52.2	197	9	US-10-180-557-448	Sequence 448, App
40	36	52.2	197	9	US-09-931-836-11	Sequence 11, Appl
41	36	52.2	197	9	US-10-173-700-448	Sequence 448, App
42	36	52.2	197	9	US-10-174-572-448	Sequence 448, App
43	36	52.2	197	9	US-10-174-579-448	Sequence 448, App
44	36	52.2	197	9	US-10-174-582-448	Sequence 448, App
45	36	52.2	197	9	US-10-174-588-448	Sequence 448, App

## ALIGNMENTS

RESULT 1  
 US-09-824-647-4  
 ; Sequence 4, Application US/09824647  
 ; Publication No. US20020183270A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Serrero, GINETTE  
 ; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
 ; FILE REFERENCE: 29996.488/P001-A  
 ; CURRENT APPLICATION NUMBER: US/09/824,647  
 ; CURRENT FILING DATE: 2001-04-04  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/991,862  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-17  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/863,862  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1997-05-23  
 ; NUMBER OF SEQ ID NOS: 17  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 4  
 ; LENGTH: 12  
 ; TYPE: PRT  
 ; ORGANISM: mouse granulins  
 ; FEATURE:  
 ; NAME/KEY: PEPTIDE  
 ; LOCATION: (1)..(12)  
 ; OTHER INFORMATION: Internal peptide of mouse GP88 used to raise the  
 ; OTHER INFORMATION: antisera against the GP88 used in the  
 ; OTHER INFORMATION: immunoaffinity step.  
 US-09-824-647-4

Query Match	100.0%;	Score 69;	DB 9;	Length 12;
Best Local Similarity	100.0%;	Pred. No. 4e-05;		
Matches 12;	Conservative:	0;	Mismatches 0;	Indels 0;
Gaps 0;				

```
Qy      1 PDAKTQCPDDST 12
        |||||
Db      1 PDAKTQCPDDST 12
```

RESULT 2  
US-10-218-509-4  
; Sequence 4, Application US/10218509  
; Publication No. US20030092661A1  
; GENERAL INFORMATION:

APPLICANT: Serrero, Ginette  
TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
FILE REFERENCE: 29996.488/P001-A  
CURRENT APPLICATION NUMBER: US/10/218,509  
CURRENT FILING DATE: 2002-08-15  
PRIOR APPLICATION NUMBER: 08/991,862  
PRIOR FILING DATE: 1998-08-17  
PRIOR APPLICATION NUMBER: 08/863,862  
PRIOR FILING DATE: 1997-05-23  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 4  
LENGTH: 12  
TYPE: PRT  
ORGANISM: mouse granulin  
FEATURE:  
NAME/KEY: PEPTIDE  
LOCATION: (1)..(12)  
OTHER INFORMATION: Internal peptide of mouse GP88 used to raise the  
OTHER INFORMATION: antisera against the GP88 used in the  
OTHER INFORMATION: immunoaffinity step.  
US-10-218-509-4

Query Match 100.0%; Score 69; DB 9; Length 12;  
Best Local Similarity 100.0%; Pred. No. 4e-05;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDAKTQCPDDST 12  
Db 1 PDAKTQCPDDST 12

## RESULT 3

US-10-281-160-4  
Sequence 4, Application US/10281160  
Publication No. US20030108950A1  
GENERAL INFORMATION:  
APPLICANT: Serrero, Ginette  
TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
FILE REFERENCE: 29996.488/P001-A  
CURRENT APPLICATION NUMBER: US/10/281,160  
CURRENT FILING DATE: 2002-10-28  
PRIOR APPLICATION NUMBER: 08/991,862  
PRIOR FILING DATE: 1998-08-17  
PRIOR APPLICATION NUMBER: 08/863,862  
PRIOR FILING DATE: 1997-05-23  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 4  
LENGTH: 12  
TYPE: PRT  
ORGANISM: mouse granulin  
FEATURE:  
NAME/KEY: PEPTIDE  
LOCATION: (1)..(12)  
OTHER INFORMATION: Internal peptide of mouse GP88 used to raise the  
OTHER INFORMATION: antisera against the GP88 used in the  
OTHER INFORMATION: immunoaffinity step.  
US-10-281-160-4

Query Match 100.0%; Score 69; DB 9; Length 12;  
Best Local Similarity 100.0%; Pred. No. 4e-05;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDAKTQCPDDST 12  
Db 1 PDAKTQCPDDST 12

## RESULT 4

US-09-813-156-4  
Sequence 4, Application US/09813156  
Patent No. US20020061859A1

GENERAL INFORMATION:  
APPLICANT: Serrero, Ginette  
TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
FILE REFERENCE: 29996.488/P001-A  
CURRENT APPLICATION NUMBER: US/09/813,156  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 08/991,862  
PRIOR FILING DATE: 1997-12-16  
PRIOR APPLICATION NUMBER: 08/863,862  
PRIOR FILING DATE: 1997-05-23  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 4  
LENGTH: 12  
TYPE: PRT  
ORGANISM: mouse granulin  
FEATURE:  
NAME/KEY: PEPTIDE  
LOCATION: (1)..(12)  
OTHER INFORMATION: Internal peptide of mouse GP88 used to raise the  
OTHER INFORMATION: antisera against the GP88 used in the  
OTHER INFORMATION: immunoaffinity step.  
US-09-813-156-4

Query Match 100.0%; Score 69; DB 10; Length 12;  
Best Local Similarity 100.0%; Pred. No. 4e-05;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDAKTQCPDDST 12  
Db 1 PDAKTQCPDDST 12

## RESULT 5

US-09-824-807-4  
Sequence 4, Application US/09824807  
Patent No. US20020094966A1  
GENERAL INFORMATION:  
APPLICANT: Serrero, Ginette  
TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
FILE REFERENCE: 29996.488/P001-A  
CURRENT APPLICATION NUMBER: US/09/824,807  
CURRENT FILING DATE: 2001-04-04  
PRIOR APPLICATION NUMBER: 08/991,862  
PRIOR FILING DATE: 1997-12-16  
PRIOR APPLICATION NUMBER: 08/863,862  
PRIOR FILING DATE: 1997-05-23  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 4  
LENGTH: 12  
TYPE: PRT  
ORGANISM: mouse granulin  
FEATURE:  
NAME/KEY: PEPTIDE  
LOCATION: (1)..(12)  
OTHER INFORMATION: Internal peptide of mouse GP88 used to raise the  
OTHER INFORMATION: antisera against the GP88 used in the  
OTHER INFORMATION: immunoaffinity step.  
US-09-824-807-4

Query Match 100.0%; Score 69; DB 10; Length 12;  
Best Local Similarity 100.0%; Pred. No. 4e-05;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDAKTQCPDDST 12  
Db 1 PDAKTQCPDDST 12

## RESULT 6

US-09-824-647-2  
Sequence 2, Application US/09824647

```
; Publication No. US20020183270A1
; GENERAL INFORMATION:
; APPLICANT: Seriero, GINETTE
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
; FILE REFERENCE: 29996.488/P001-A
; CURRENT APPLICATION NUMBER: US/09/824.647
; CURRENT FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/991,862
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/863,862
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-05-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Mouse epithelin/granulin
US-09-824-647-2

Query Match      100.0%; Score 69; DB 9; Length 589;
Best Local Similarity 100.0%; Pred. No. 0.0021;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PDAKTQCPDDST 12
        |||||
DB      208 PDAKTQCPDDST 219

RESULT 7
US-10-218-509-2
; Sequence 2, Application US/10218509
; Publication No. US20030092661A1
; GENERAL INFORMATION:
; APPLICANT: Seriero, GINETTE
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
; FILE REFERENCE: 29996.488/P001-A
; CURRENT APPLICATION NUMBER: US/10/218.509
; CURRENT FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: 08/991,862
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 08/863,862
; PRIOR FILING DATE: 1997-05-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Mouse epithelin/granulin
US-10-218-509-2

Query Match      100.0%; Score 69; DB 9; Length 589;
Best Local Similarity 100.0%; Pred. No. 0.0021;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PDAKTQCPDDST 12
        |||||
DB      208 PDAKTQCPDDST 219

RESULT 8
US-10-281-160-2
; Sequence 2, Application US/10281160
; Publication No. US20030108950A1
; GENERAL INFORMATION:
; APPLICANT: Seriero, GINETTE
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
; FILE REFERENCE: 29996.488/P001-A
; CURRENT APPLICATION NUMBER: US/10/281.160
; CURRENT FILING DATE: 2002-10-28
; PRIOR APPLICATION NUMBER: US/08/991,862
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 08/863,862
; PRIOR FILING DATE: 1997-05-23
```

```
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Mouse epithelin/granulin
US-10-281-160-2

Query Match      100.0%; Score 69; DB 9; Length 589;
Best Local Similarity 100.0%; Pred. No. 0.0021;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PDAKTQCPDDST 12
        |||||
DB      208 PDAKTQCPDDST 219

RESULT 9
US-09-813-156-2
; Sequence 2, Application US/09813156
; Patent No. US20020061859A1
; GENERAL INFORMATION:
; APPLICANT: Seriero, GINETTE
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
; FILE REFERENCE: 29996.488/P001-A
; CURRENT APPLICATION NUMBER: US/09/813.156
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 08/991,862
; PRIOR FILING DATE: 1997-12-16
; PRIOR APPLICATION NUMBER: 08/863,862
; PRIOR FILING DATE: 1997-05-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Mouse epithelin/granulin
US-09-813-156-2

Query Match      100.0%; Score 69; DB 10; Length 589;
Best Local Similarity 100.0%; Pred. No. 0.0021;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PDAKTQCPDDST 12
        |||||
DB      208 PDAKTQCPDDST 219

RESULT 10
US-09-824-807-2
; Sequence 2, Application US/09824807
; Patent No. US20020094966A1
; GENERAL INFORMATION:
; APPLICANT: Seriero, GINETTE
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
; FILE REFERENCE: 29996.488/P001-A
; CURRENT APPLICATION NUMBER: US/09/824.807
; CURRENT FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 08/991,862
; PRIOR FILING DATE: 1997-12-16
; PRIOR APPLICATION NUMBER: 08/863,862
; PRIOR FILING DATE: 1997-05-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Mouse epithelin/granulin
US-09-824-807-2

Query Match      100.0%; Score 69; DB 10; Length 589;
Best Local Similarity 100.0%; Pred. No. 0.0021;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PDAKTQCPDDST 12
        |||||
DB      208 PDAKTQCPDDST 219
```



QY 1 PDAKTQCPDDST 12  
 |||||  
 Db 208 PDAKTQCPDDST 219

## RESULT 11

US-09-824-647-17  
 ; Sequence 17, Application US/09824647  
 ; Publication No. US20020183270A1  
 ; GENERAL INFORMATION:

; APPLICANT: Serrero, Ginette  
 ; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
 ; FILE REFERENCE: Z9996.488/P001-A  
 ; CURRENT APPLICATION NUMBER: US/09/824,647  
 ; CURRENT FILING DATE: 2001-04-04  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/991,862  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-17  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/863,862  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1997-05-23  
 ; NUMBER OF SEQ ID NOS: 17  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 17  
 ; LENGTH: 593  
 ; TYPE: PRT  
 ; ORGANISM: Human GP88 cdNA  
 US-09-824-647-17

Query Match 73.9%; Score 51; DB 9; Length 593;  
 Best Local Similarity 66.7%; Pred. No. 1.8;  
 Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 PDAKTQCPDDST 12  
 |||||  
 Db 209 PDAKRCPDGST 220

## RESULT 12

US-10-218-509-17  
 ; Sequence 17, Application US/10218509  
 ; Publication No. US20030092661A1  
 ; GENERAL INFORMATION:

; APPLICANT: Serrero, Ginette  
 ; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
 ; FILE REFERENCE: Z9996.488/P001-A  
 ; CURRENT APPLICATION NUMBER: US/10/218,509  
 ; CURRENT FILING DATE: 2002-08-15  
 ; PRIOR APPLICATION NUMBER: 08/991,862  
 ; PRIOR FILING DATE: 1998-08-17  
 ; PRIOR APPLICATION NUMBER: 08/863,862  
 ; PRIOR FILING DATE: 1997-05-23  
 ; NUMBER OF SEQ ID NOS: 17  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 17  
 ; LENGTH: 593  
 ; TYPE: PRT  
 ; ORGANISM: Human GP88 cdNA  
 US-10-218-509-17

Query Match 73.9%; Score 51; DB 9; Length 593;  
 Best Local Similarity 66.7%; Pred. No. 1.8;  
 Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 PDAKTQCPDDST 12  
 |||||  
 Db 209 PDAKRCPDGST 220

## RESULT 13

US-10-281-160-17  
 ; Sequence 17, Application US/10281160  
 ; Publication No. US20030108950A1  
 ; GENERAL INFORMATION:

; APPLICANT: Serrero, Ginette  
 ; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
 ; FILE REFERENCE: Z9996.488/P001-A  
 ; CURRENT APPLICATION NUMBER: US/10/281,160  
 ; CURRENT FILING DATE: 2002-10-28  
 ; PRIOR APPLICATION NUMBER: US/08/991,862  
 ; PRIOR FILING DATE: 1998-08-17  
 ; PRIOR APPLICATION NUMBER: 08/863,862  
 ; PRIOR FILING DATE: 1997-05-23  
 ; NUMBER OF SEQ ID NOS: 17  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 17  
 ; LENGTH: 593  
 ; TYPE: PRT  
 ; ORGANISM: Human GP88 cdNA  
 US-10-281-160-17

Query Match 73.9%; Score 51; DB 9; Length 593;  
 Best Local Similarity 66.7%; Pred. No. 1.8;  
 Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 PDAKTQCPDDST 12  
 |||||  
 Db 209 PDAKRCPDGST 220

## RESULT 14

US-09-813-156-17  
 ; Sequence 17, Application US/09813156  
 ; Patent No. US20020061859A1  
 ; GENERAL INFORMATION:

; APPLICANT: Serrero, Ginette  
 ; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
 ; FILE REFERENCE: Z9996.488/P001-A  
 ; CURRENT APPLICATION NUMBER: US/09/813,156  
 ; CURRENT FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: 08/991,862  
 ; PRIOR FILING DATE: 1997-12-16  
 ; PRIOR APPLICATION NUMBER: 08/863,862  
 ; PRIOR FILING DATE: 1997-05-23  
 ; NUMBER OF SEQ ID NOS: 17  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 17  
 ; LENGTH: 593  
 ; TYPE: PRT  
 ; ORGANISM: Human GP88 cdNA  
 US-09-813-156-17

Query Match 73.9%; Score 51; DB 10; Length 593;  
 Best Local Similarity 66.7%; Pred. No. 1.8;  
 Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 PDAKTQCPDDST 12  
 |||||  
 Db 209 PDAKRCPDGST 220

## RESULT 15

US-09-824-807-17  
 ; Sequence 17, Application US/09824807  
 ; Patent No. US20020094966A1  
 ; GENERAL INFORMATION:

; APPLICANT: Serrero, Ginette  
 ; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
 ; FILE REFERENCE: Z9996.488/P001-A  
 ; CURRENT APPLICATION NUMBER: US/09/824,807  
 ; CURRENT FILING DATE: 2001-04-04  
 ; PRIOR APPLICATION NUMBER: 08/991,862  
 ; PRIOR FILING DATE: 1997-12-16  
 ; PRIOR APPLICATION NUMBER: 08/863,862  
 ; PRIOR FILING DATE: 1997-05-23  
 ; NUMBER OF SEQ ID NOS: 17  
 ; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 17  
; LENGTH: 593  
; TYPE: PRT  
; ORGANISM: Human GP88 cDNA  
US-09-824-807-17

Query Match 73.9%; Score 51; DB 10; Length 593;  
Best Local Similarity 66.7%; Pred. No. 1.8;  
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 PDAKTQCPDDST 12  
| | | | | | | | | |  
Db 209 PDARSRCPDGST 220

Search completed: July 8, 2003, 16:31:40  
Job time : 12.6923 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2003, 16:19:45 ; Search time 12.1538 seconds  
(without alignments)  
94.918 Million cell updates/sec

Title: US-09-824-647-4  
Perfect score: 69  
Sequence: 1 PDAKTCPPDDST 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_73:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	69	100.0	589	2 C38128	epithelin/granulin
2	69	100.0	589	2 B38128	epithelin/granulin
3	56	81.2	591	2 I48141	acroganin - guine
4	51	73.9	593	1 GIHU	granulin precursor
5	43	62.3	318	2 E87929	protein T2H2.6 [1
6	43	62.3	345	2 T25138	hypothetical prote
7	43	62.3	358	2 T25137	hypothetical prote
8	42	60.9	506	2 AC3545	periplasmic dipept
9	40	58.0	1423	2 A49206	exo-beta-D-fructos
10	39.5	57.2	234	1 JX0197	ribonuclease TRV (
11	39	56.5	190	2 T33366	hypothetical prote
12	39	56.5	222	2 T16381	hypothetical prote
13	39	56.5	465	2 H97712	hypothetical prote
14	38	55.1	57	2 A46654	growth modulatory
15	38	55.1	144	2 C84211	iron-dependent rep
16	38	55.1	351	2 AG1436	transcription regu
17	38	55.1	351	2 AH1078	transcription regu
18	38	55.1	436	2 E69371	bile acid-inducibl
19	38	55.1	479	2 S18447	variant surface gl
20	38	55.1	706	2 F87683	peptidase M13 fami
21	38	55.1	1283	2 T28812	hypothetical prote
22	37	53.6	127	2 AH2701	hypothetical prote
23	37	53.6	169	2 T32698	hypothetical prote
24	37	53.6	186	2 T35736	hypothetical prote
25	37	53.6	249	2 C96691	probable wrky-type
26	37	53.6	318	2 A29636	myoblast determina
27	37	53.6	340	2 A87388	transcription regu
28	37	53.6	381	2 H83985	alpha-D-mannose-al
29	37	53.6	427	2 AH0728	probable membrane

30 37 53.6 583 2 G84829  
31 37 53.6 601 2 T18787  
32 37 53.6 1249 2 T14150  
33 37 53.6 1441 2 B86807  
34 37 53.6 2090 2 S26058  
35 37 53.6 2960 2 A45259  
36 37 53.6 3566 1 A40701  
37 36 52.2 181 2 T32690  
38 36 52.2 245 2 T33363  
39 36 52.2 289 2 AG3260  
40 36 52.2 319 2 S26827  
41 36 52.2 346 2 T33365  
42 36 52.2 357 2 T23460  
43 36 52.2 419 2 F96704  
44 36 52.2 447 2 C64243  
45 36 52.2 456 2 AB2425

probable. pIR2 fami  
hypothetical prote  
vesicle associated  
hypothetical prote  
probable transfor  
desmoyokin - human  
tenascin-X precurs  
hypothetical prote  
hypothetical prote  
exodeoxyribonuclea  
myogenic factor 3  
hypothetical prote  
hypothetical prote  
aminopeptidase (EC  
serine/threonine k

## ALIGNMENTS

### RESULT 1

C38128  
epithelin/granulin precursor - mouse  
N: Alternate names: acroganin; PC-cell-derived growth factor  
C: Species: Mus musculus (house mouse)  
C: Date: 10-Jul-1992 #sequence\_revision 10-Jul-1992 #text\_change 20-Aug-1999  
C: Accession: C38128; S32503; I49468; A46705  
R: Plowman, G.D.; Green, J.M.; Neubauer, M.G.; Buckley, S.D.; McDonald, V.L.; Tod  
J. Biol. Chem. 267, 13073-13078, 1992  
A: Title: The epithelin precursor encodes two proteins with opposing activities o  
A: Reference number: A38128; MUID:92317004; PMID:1618805  
A: Accession: C38128  
A: Molecule type: mRNA  
A: Residues: 1-589 <P>  
A: Cross-references: GB:X63221; NID:950851; PIDN:CAA4197.1; PID:950852  
R: Baba, T.; Nemoto, H.; Watanabe, K.; Arai, Y.; Gerton, G.L.  
FEBS Lett. 322, 89-94, 1993  
A: Title: Exon/intron organization of the gene encoding the mouse epithelin/granu  
A: Reference number: S32503; MUID:93245991; PMID:8482392  
A: Accession: S32503  
A: Molecule type: DNA  
A: Residues: 18-349, 'L', 351-589 <BAB>  
R: Baba, T.; Hoff, H.B.  
Mol. Reprod. Dev. 34, 233-243, 1993  
A: Title: Acroganin, an acrosomal cysteine-rich glycoprotein, is the precursor of  
A: Reference number: I48141; MUID:93228994; PMID:8471244  
A: Accession: I49468  
A: Status: preliminary; translated from GB/EMBL/DBJ  
A: Molecule type: mRNA  
A: Residues: 1-250, 'L', 252-253, 'V', 255-349, 'L', 351-401, 'SA', 404-589 <RES>  
A: Cross-references: GB:M86736; NID:g191766; PIDN:AAA37191.1; PID:g191767  
R: Zhou, J.; Gao, G.; Crabb, J.W.; Serreto, G.  
J. Biol. Chem. 268, 10863-10869, 1993  
A: Title: Purification of an autocrine growth factor homologous with mouse epithe  
A: Reference number: A46705; MUID:93266526; PMID:8496151  
A: Accession: A46705  
A: Status: preliminary  
A: Molecule type: protein  
A: Residues: 18-19, 'X', 21-25, 'X', 27-29, 'XX', 32, 'XXX', 119-127, 152-154, 'DXK', 158-16  
C: Superfamily: granulin

Query Match 100.0%; Score 69; DB 2; Length 589;

.Best Local Similarity 100.0%; Pred. No. 0.0011;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PDAKTCPPDDST 12

Db 208 PDAKTCPPDDST 219

### RESULT 2

B38128

```
Qy      1 PDAKTQCpDDST 12
        ||::|||:|
Db     209 PDSRSQCpDDTT 220
```

F;18-593/Product: granulin #98

```
Qy      1 PDAKTQCpDDST 12
        ||::|||:|
Db     209 PDSRSQCpDDTT 220
```

F:18-44/Product: paranulin #status: experimental <PGR>  
 F:58-113/Product: granulin G #status: predicted <GRG>  
 F:123-179/Product: granulin F #status: predicted <GRF>  
 F:206-261/Product: granulin B #status: experimental <GRB>  
 F:281-336/Product: granulin A #status: experimental <GRA>  
 F:364-417/Product: granulin C #status: experimental <GRC>  
 F:442-496/Product: granulin D #status: predicted <GRD>  
 F:518-573/Product: granulin E #status: predicted <GRE>  
 F:368/Blinding site: carbohydrate (asn) (covalent) #status: experimental

Query Match 73.9%; Score 51; DB 1; Length 593;  
 Best Local Similarity 66.7%; Pred. No. 0.97; Mismatches 1; Indels 0; Gaps 0;  
 Matches 8; Conservative 3;

QY 1 PDATQCPDDST 12  
 ||:|:| ||  
 Db 209 PDARSRCPDGST 220

## RESULT 5

E87929 protein T22H2.6 [Imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans  
 C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
 C:Accession: E87929  
 R:Anonymous, The C. elegans Sequencing Consortium.  
 Science 282, 2012-2018, 1998  
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biological processes.  
 A:Reference number: A75000; PMID:99069613; PMID:9851916  
 A:Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_elegans/  
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and Science 283, 2103, 1999.  
 A:Accession: E87929  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-318 <STO>  
 A:Cross-references: GB:chr\_I; PIDN:CAB04752.1; PID:G3880056; GSPDB:GN00019; CESP:T22H2.6  
 C:Genetics:  
 A:Gene: T22H2.6  
 A:Map position: 1

Query Match 62.3%; Score 43; DB 2; Length 318;  
 Best Local Similarity 63.6%; Pred. No. 11; Mismatches 2; Indels 0; Gaps 0;  
 Matches 7; Conservative 2;

QY 2 DAKTQCPDDST 12  
 ||:|:| ||  
 Db 59 DAETECSDDET 69

## RESULT 6

T25138 hypothetical protein T22H2.6b - Caenorhabditis elegans

C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T25138  
 R:Lennard, N.  
 submitted to the EMBL Data Library, November 1996  
 A:Reference number: Z19985  
 A:Accession: T25138  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-345 <WIL>  
 A:Cross-references: EMBL:Z81595; PIDN:CAB54305.1; GSPDB:GN00019; CESP:T22H2.6b  
 A:Experimental source: clone T22H2  
 C:Genetics:  
 A:Gene: CESP:T22H2.6b  
 A:Map position: 1  
 A:Introns: 93/3; 232/3; 314/3

Query Match 62.3%; Score 43; DB 2; Length 345;  
 Best Local Similarity 63.6%; Pred. No. 12; Mismatches 2; Indels 0; Gaps 0;  
 Matches 7; Conservative 2;

QY 2 DAKTQCPDDST 12  
 ||:|:| ||  
 Db 99 DAETECSDDET 109

## RESULT 7

T25137 hypothetical protein T22H2.6a - Caenorhabditis elegans

C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T25137  
 R:Lennard, N.  
 submitted to the EMBL Data Library, November 1996  
 A:Reference number: Z19985  
 A:Accession: T25137  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-358 <WIL>  
 A:Cross-references: EMBL:Z81595; PIDN:CAB54304.1; GSPDB:GN00019; CESP:T22H2.6a  
 A:Experimental source: clone T22H2  
 C:Genetics:  
 A:Gene: CESP:T22H2.6a  
 A:Map position: 1  
 A:Introns: 93/3; 232/3; 314/3

Query Match 62.3%; Score 43; DB 2; Length 358;  
 Best Local Similarity 63.6%; Pred. No. 13; Mismatches 2; Indels 0; Gaps 0;  
 Matches 7; Conservative 2;

QY 2 DAKTQCPDDST 12  
 ||:|:| ||  
 Db 99 DAETECSDDET 109

## RESULT 8

AC3545 periplasmic dipeptide transport protein precursor BMEII0284 [Imported] - Brucella

C:Species: Brucella melitensis  
 C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 01-Feb-2002  
 C:Accession: AC3545  
 R:DeIvecchio, V.G.; Kaputal, V.; Redkar, R.J.; Patra, G.; Mijer, C.; Los, T.; I.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
 A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis.  
 A:Reference number: AD3252; PMID:11756688  
 A:Accession: AC3545  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-506 <KUR>  
 A:Cross-references: GB:AE008918; PIDN:AAL53526.1; PID:G17984432; GSPDB:GN00191  
 A:Experimental source: strain 16M  
 C:Genetics:  
 A:Gene: BMEII0284  
 A:Map position: 11

Query Match 60.9%; Score 42; DB 2; Length 506;  
 Best Local Similarity 58.3%; Pred. No. 25; Mismatches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 PDATQCPDDST 12  
 ||:|:| ||  
 Db 143 PIATCECPDPAT 154

## RESULT 9

A49206 exo-beta-D-fructosidase - Streptococcus mutans

C:Species: Streptococcus mutans  
 C:Date: 19-Dec-1993 #sequence\_revision 18-Nov-1994 #text\_change 21-Jul-2000  
 C:Accession: A49206  
 R:Burne, R.A.; Penders, J.E.; Infect. Immun. 60, 4621-4632, 1992  
 A:Title: Characterization of the Streptococcus mutans GS-5 frua gene encoding ex-

A:Reference number: A49206; MUID:93014178; PMID:1398976

A:Contents: GS-5

A:Accession: A49206

A>Status: Preliminary

A:Molecule type: nucleic acid

A:Residues: 1-1423 <BUT>

A:Cross-references: GB:L03358; NID:94733888; PIDN:AAA26889.1; PID:gl53634

A>Note: sequence extracted from NCBI backbone (NCBIN:116439, NCBIP:116440)

Query Match 58.0%; Score 40; DB 2; Length 1423;

Best Local Similarity 66.7%; Pred. No. 1.4e+02;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PDAKTQCPD 9

|| |::|||

Db 619 PDLHTECPD 627

#### RESULT 10

JX0197

Ribonuclease TRV (EC 3.1.27.-) - fungus (*Trichoderma viride*)

C:Species: *Trichoderma viride*

C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999

C:Accession: JX0197

R:Inada, Y.; Watanabe, H.; Ohgii, K.; Irie, M.

J. Biochem. 110, 896-904, 1991

A:Reference number: JX0197; MUID:92176163; PMID:1794979

A:Accession: JX0197

A:Molecule type: protein

A:Residues: 1-234 <INA>

C:Superfamily: Enterobacter ribonuclease

C:Keywords: glycoprotein; hydrolase

F:15/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 57.2%; Score 39.5; DB 1; Length 234;

Best Local Similarity 72.7%; Pred. No. 32;

Matches 8; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 PD-AKTQCPD 10

|| |::|||

Db 217 PDGAKSTCPD 227

#### RESULT 11

T33366

hypothetical protein F16G10.1 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 20-Jun-2000

C:Accession: T33366

R:Gattung, S.; Scheet, P.

submitted to the EMBL Data Library, July 1998

A:Description: The sequence of *C. elegans* cosmid F16G10.

A:Reference number: Z21329

A:Accession: T33366

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-190 <GAT>

A:Cross-references: EMBL:AF077537; PIDN:AAC26282.1; GSPDB:GN00020; CESP:F16G10.1

A:Experimental source: strain Bristol N2; clone F16G10

C:Genetics:

A:Gene: CESP:F16G10.1

A:Map position: 2

A:Introns: 61/3; 84/1; 127/3

C:Superfamily: *Caenorhabditis elegans* hypothetical protein F16G10.4

Query Match 56.5%; Score 39; DB 2; Length 190;

Best Local Similarity 58.3%; Pred. No. 32;

Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 PDAKTQCPDST 12

|:|::|||

Db 69 PTSTTKCPDLST 80

#### RESULT 12

T16381

hypothetical protein F47B7.4 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Jun-2000

C:Accession: T16381

R:Leimbach, D.

submitted to the EMBL Data Library, October 1995

A:Description: The sequence of *C. elegans* cosmid F47B7.

A:Reference number: Z18504

A:Accession: T16381

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-222 <LEI>

A:Cross-references: EMBL:U39646; NID:91049351; PID:91049353; PIDN:AAA80368.1; CESP:

C:Genetics:

A:Gene: CESP:F47B7.4

A:Introns: 81/3; 97/1; 116/1; 163/3; 212/1

C:Superfamily: *Caenorhabditis elegans* hypothetical protein F16G10.4

Query Match 56.5%; Score 39; DB 2; Length 222;

Best Local Similarity 58.3%; Pred. No. 37;

Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 PDAKTQCPDST 12

|:|::|||

Db 101 PTSTTKCPDLST 112

#### RESULT 13

H97712

hypothetical protein pntB [imported] - *Rickettsia conorii* (strain Malish 7)

C:Species: *Rickettsia conorii*

C>Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 22-Oct-2001

C:Accession: H97712

R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, I.

Science 293, 2093-2098, 2001

A:Title: Mechanisms of Evolution in *Rickettsia conorii* and *Rickettsia prowazekii*.

A:Reference number: A97700; MUID:21442074; PMID:11557893

A:Accession: H97712

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-465 <KUR>

A:Cross-references: GB:AE006914; PIDN:AAL02642.1; PID:gl5619144; GSPDB:GN00173

C:Genetics:

A:Gene: pntB

C:Superfamily: NAD(P)+ transhydrogenase (B-specific) beta chain; NAD(P)+ transhydr

Query Match 56.5%; Score 39; DB 2; Length 465;

Best Local Similarity 66.7%; Pred. No. 72;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 AKTQCPDSS 11

|||::|||

Db 291 AKTSCPEDA 299

#### RESULT 14

A46654

growth modulatory factor granulin-1 - common carp

C:Species: *Cyprinus carpio* (common carp)

C>Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 10-Nov-1995

C:Accession: A46654; A40180

R:Belcourt, D.R.; Lazure, C.; Bennett, H.P.

J. Biol. Chem. 268, 9230-9237, 1993

A:Title: Isolation and primary structure of the three major forms of granulin-like

A:Reference number: A46654; MUID:93252781; PMID:8486624

A:Accession: A46654

A:Molecule type: protein

A:Residues: 1-57 <BEL>

A:Experimental source: spleen and head kidney

A:Note: sequence extracted from NCBI backbone (NCBIP:131314)  
C:Comment: All twelve Cys residues are involved in disulfide bonds.  
C:Keywords: disulfide bond; monomer

Query Match 55.1%; Score 38; DB 2; Length 57;  
Best Local Similarity 63.6%; Pred. No. 16;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 DAKTOCPDST 12  
|||  
Db 5 DAKTICPGTT 15

## RESULT 15

C84211  
Iron-dependent repressor [Imported] - Halobacterium sp. NRC-1  
C:Species: Halobacterium sp. NRC-1  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C:Accession: C84211  
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.;  
Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablon  
Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li  
A:Title: Genome sequence of Halobacterium species NRC-1.  
A:Reference number: A84160; MUID:20504483; PMID:11016950  
A:Accession: C84211  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-144 <STO>  
A:Cross-references: GB:AB004437; NID:g10580129; PIDN:AAG19055.1; GSPDB:GN00138  
C:Genetics:  
A:Gene: troR

Query Match 55.1%; Score 38; DB 2; Length 144;  
Best Local Similarity 75.0%; Pred. No. 36;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 DAKTOCPD 9  
|||  
Db 116 DRKSQCPD 123

Search completed: July 8, 2003, 16:29:25  
Job time : 15.1538 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2003, 16:16:30 ; Search time 5.84615 Seconds  
(without alignments)  
85.136 Million cell updates/sec

Title: US-09-824-647-4

Perfect score: 69

Sequence: 1 PDAKTCQPDST 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	69	100.0	588	1 GRN_RAT	P23785 r granulins
2	69	100.0	589	1 GRN_MOUSE	P28798 mus musculus
3	56	81.2	591	1 GRN_CAVPO	P28797 cavia porce
4	51	73.9	593	1 GRN_HUMAN	P28799 h granulins
5	41	59.4	564	1 CDC7_MOUSE	Q920h0 mus musculus
6	40	58.0	1423	1 FRC4_STRMU	Q03174 streptococ
7	39.5	57.2	234	1 RNRTR_TRIVI	P24657 trichoderna
8	38	55.1	57	1 GRN1_CYPCA	P81013 cyprinus ca
9	38	55.1	479	1 VS12_TRYBB	P26327 trypanosoma
10	38	55.1	511	1 VGIG_VSVO	P04884 vesicular s
11	38	55.1	993	1 VIA_CWVII	Q83270 cucumber mo
12	37	53.6	318	1 MYOD_MOUSE	P10085 mus musculus
13	37	53.6	2090	1 N214_HUMAN	P35658 homo sapien
14	37	53.6	2960	1 AHNK_HUMAN	Q09666 homo sapien
15	37	53.6	4289	1 TENX_HUMAN	P22105 homo sapien
16	36	52.2	116	1 SELX_MOUSE	Q9j1c3 mus musculus
17	36	52.2	197	1 I17C_HUMAN	Q9p0m4 homo sapien
18	36	52.2	319	1 MYOD_HUMAN	P15172 homo sapien
19	36	52.2	319	1 MYOD_FIG	P49811 sus scrofa
20	36	52.2	447	1 AMPA_MYCGE	P47631 mycoplasma
21	36	52.2	452	1 GUX1_CRYPA	Q00548 cryptoneutr
22	36	52.2	574	1 CDC7_HUMAN	O00311 homo sapien
23	36	52.2	627	1 NDF8_YEAST	P38830 saccharomyc
24	36	52.2	1121	1 ALP1_SCHPO	Q10197 schizosacch
25	36	52.2	1480	1 SLIT_DROME	P24014 drosophila
26	35.5	51.4	204	1 YIGZ_ECOLI	P27862 escherichia
27	35.5	51.4	489	1 S0F1_YEAST	P33750 saccharomyc
28	35	50.7	107	1 Y866_RICPR	Q9zca0 rickettsia
29	35	50.7	369	1 PAR3_MOUSE	Q08675 mus musculus
30	35	50.7	372	1 LI66_PHACH	P50622 phanerocae
31	35	50.7	388	1 P2X4_HUMAN	Q99571 homo sapien
32	35	50.7	757	1 COMP_HUMAN	P49747 homo sapien
33	35	50.7	809	1 UBP1_YEAST	P25037 saccharomyc

34	35	50.7	991	1	DHPL_SCHPO	P40848 schizosacch
35	35	50.7	1002	1	YEMA_DROME	P25992 drosophila
36	35	50.7	1738	1	CO4_MOUSE	P01029 mus musculus
37	35	50.7	1877	1	PK5_MOUSE	Q04592 mus musculus
38	34	49.3	60	1	TXW1_OPHHA	P01412 ophiophagus
39	34	49.3	81	1	NX54_AIPLA	P19960 aipysurus l
40	34	49.3	96	1	FER_SCEQU	P00238 scenedesmus
41	34	49.3	98	1	FER1_CYAPA	P17007 cyanophora
42	34	49.3	110	1	RPCZ_YEAST	Q04307 saccharomyc
43	34	49.3	155	1	NRDG_VIBCH	Q9km76 vibrio chol
44	34	49.3	166	1	Y443_HAEIN	P44101 haemophilus
45	34	49.3	244	1	VC23_VACCC	P21090 vaccinia v1

## ALIGNMENTS

### RESULT 1

GRN_RAT	GRN_RAT	STANDARD	PRT	588 AA
AC	P23785			
DT	01-NOV-1991 (Rel. 20, Created)			
DT	01-DEC-1992 (Rel. 24, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Granulins precursor (Acrogranin) [Contains: Granulin 1 (Granulin G); Granulin 2 (Granulin F); Granulin 3 (Granulin B) (Epithelin 2); Granulin 4 (Granulin A) (Epithelin 1); Granulin 5 (Granulin C); Granulin 6 (Granulin D); Granulin 7 (Granulin E)].			
GN	GRN.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Kidney;			
RA	MEDLINE=94062640; PubMed=8243292;			
RA	Bhandari V., Glaid A., Bateman A.;			
RT	"The complementary deoxyribonucleic acid sequence, tissue distribution, and cellular localization of the rat granulin precursor."			
RT	Endocrinology 133:2682-2689(1993).			
RL	[2]			
RN	SEQUENCE FROM N.A., AND SEQUENCE OF 204-259 AND 278-334.			
RP	TISSUE=Kidney;			
RC	MEDLINE=92317004; PubMed=1618805;			
RA	Plowman G.D., Green J.M., Neubauer M.G., Buckley S.D., McDonald V.L.,			
RA	Todaro G.J., Shoyab M.;			
RT	"The epithelin precursor encodes two proteins with opposing activities on epithelial cell growth."			
RL	J. Biol. Chem. 267:13073-13078(1992).			
RN	[3]			
RP	SEQUENCE OF 204-225 AND 279-299.			
RC	MEDLINE=91045907; PubMed=2236009;			
RA	Shoyab M., McDonald V.L., Byles C., Todaro G.J., Plowman G.D.;			
RT	"Epithelins 1 and 2: isolation and characterization of two cysteine-rich growth-modulating proteins."			
RL	Proc. Natl. Acad. Sci. U.S.A. 87:7912-7916(1990).			
RN	[4]			
RP	SEQUENCE OF 278-328.			
RC	TISSUE=Bone marrow;			
RA	MEDLINE=91097544; PubMed=2268320;			
RA	Bateman A., Belcourt D.R., Bennett H.P., Lazure C., Solomon S.;			
RT	"Granulins, a novel class of peptide from leukocytes."			
RL	Biochem. Biophys. Res. Commun. 173:1161-1168(1990).			
CC	-I- FUNCTION: GRANULINS HAVE POSSIBLE CYTOKINE-LIKE ACTIVITY. THEY MAY PLAY A ROLE IN INFLAMMATION, WOUND REPAIR, AND TISSUE REMODELING.			
CC	-I- TISSUE SPECIFICITY: UBIQUITOUS; MOST ABUNDANT IN THE SPLEEN AND SEVERAL TISSUES OF ENDOCRINE SIGNIFICANCE.			
CC	-I- PTM: GRANULINS ARE DISULFIDE BRIDGED.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -			



CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; M97750; AAL16903.1; -  
 CC EMBL; X62322; CAA44198.1; -  
 CC PIR; A36199; A36199.  
 CC PIR; B36199; B36199.  
 CC PIR; E36698; E36698.  
 CC InterPro; IPR000118; Granulin.  
 CC Pfam; PF00396; granulin; 7.  
 CC SMART; SM00799; GRANULIN; 7.  
 CC PROSITE; PS00799; GRANULINS; 7.  
 CC CYTOKINE; Repeat; Glycoprotein; Signal.  
 CC SIGNAL 1 17 POTENTIAL.  
 CC CHAIN 18 588 ACROGRANIN.  
 CC PEPTIDE 58 113 GRANULIN 1.  
 CC PEPTIDE 122 178 GRANULIN 2.  
 CC PEPTIDE 204 259 GRANULIN 3.  
 CC PEPTIDE 278 334 GRANULIN 4.  
 CC PEPTIDE 361 413 GRANULIN 5.  
 CC PEPTIDE 438 492 GRANULIN 6.  
 CC PEPTIDE 512 567 GRANULIN 7.  
 CC CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 372 372 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 525 525 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CONFLICT 201 201 S -> FP (IN REF. 2).  
 CC CONFLICT 307 308 TK -> SB (IN REF. 4).  
 CC CONFLICT 324 324 Q -> T (IN REF. 4).  
 CC CONFLICT 388 388 M -> I (IN REF. 2).  
 CC SEQUENCE 588 AA; 63369 MW; 113D434F7E099B31 CRC64;

Query Match 100.0%; Score 69; DB 1; Length 588;  
 Best Local Similarity 100.0%; Pred. No. 0.00043;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDAKTQCPDST 12  
 |||||  
 DB 207 PDAKTQCPDST 218

RESULT 2  
 GRN\_MOUSE STANDARD; PRT; 589 AA.  
 AC P28798;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Granulins precursor (Acrogranin) [Contains: Granulin 1; Granulin 2;  
 DE Granulin 3; Granulin 4; Granulin 5; Granulin 6; Granulin 7].  
 GN GRN.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93245991; PubMed=8482392;  
 RA Baba T., Nemoto H., Watanabe K., Arai Y., Gerton G.L.;  
 RT "Exon/intron organization of the gene encoding the mouse  
 RT epithelin/granulin precursor (acrogranin).";  
 RL FEBS Lett. 322:89-94(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=92317004; PubMed=1618805;  
 RA Plozman G.D., Green I.M., Neubauer M.G., Buckley S., McDonald V.L.,  
 RA Todaro G.I., Shoyab M.;  
 RT "The epithelin precursor encodes two proteins with opposing  
 RT activities on epithelial cell growth.";  
 RL J. Biol. Chem. 267:13073-13078(1992).

CC -!- FUNCTION: GRANULINS HAVE POSSIBLE CYTOKINE-LIKE ACTIVITY. THEY MAY  
 CC PLAY A ROLE IN INFLAMMATION, WOUND REPAIR, AND TISSUE REMODELING.  
 CC -!- TISSUE SPECIFICITY: UBICUITOUS.  
 CC -!- PTM: GRANULINS ARE DISULFIDE BRIDGED.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; D16195; BAA03736.1; -  
 CC EMBL; M86736; AAA37191.1; -  
 CC EMBL; X62321; CAA44197.1; -  
 CC MGD; MGI:95832; Grn.  
 CC InterPro; IPR000118; Granulin.  
 CC Pfam; PF00396; granulin; 7.  
 CC SMART; SM00799; GRANULINS; 7.  
 CC PROSITE; PS00799; GRANULINS; 7.  
 CC CYTOKINE; Repeat; Glycoprotein; Signal.  
 CC SIGNAL 1 17 POTENTIAL.  
 CC CHAIN 18 589 ACROGRANIN.  
 CC PEPTIDE 758 7113 GRANULIN 1.  
 CC PEPTIDE 7122 7178 GRANULIN 2.  
 CC PEPTIDE 205 260 GRANULIN 3.  
 CC PEPTIDE 280 334 GRANULIN 4.  
 CC PEPTIDE 362 7414 GRANULIN 5.  
 CC PEPTIDE 440 7493 GRANULIN 6.  
 CC PEPTIDE 7517 7568 GRANULIN 7.  
 CC CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 373 373 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 526 526 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CONFLICT 350 350 L -> R (IN REF. 2).  
 CC SEQUENCE 589 AA; 63458 MW; 1DE8229C413CB787 CRC64;

Query Match 100.0%; Score 69; DB 1; Length 589;  
 Best Local Similarity 100.0%; Pred. No. 0.00043;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDAKTQCPDST 12  
 |||||  
 DB 208 PDAKTQCPDST 219

RESULT 3  
 GRN\_CAVPO STANDARD; PRT; 591 AA.  
 AC P28797;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Granulins precursor (Acrogranin) [Contains: Granulin 1; Granulin 2;  
 DE Granulin 3; Granulin 4; Granulin 5; Granulin 6; Granulin 7]  
 DE (Fragment).  
 GN GRN.  
 OS Cavia porcellus (Guinea pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Mystricognathi; Caviidae; Cavia.  
 OX NCBI\_TaxID=10141;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 4-21.  
 RC TISSUE=Testis;  
 RX MEDLINE=93228994; PubMed=8471244;  
 RA Baba T., Hoff H.B. III, Nemoto H., Lee H., Orth J., Arai Y.,  
 RA Gerton G.L.;  
 RA "Acrogranin, an acrosomal cysteine-rich glycoprotein, is the  
 RA precursor of the growth-modulating peptides, granulins, and  
 RA epithelins, and is expressed in somatic as well as male germ cells.";  
 RL Mol. Reprod. Dev. 34:233-243(1993).  
 CC -!- FUNCTION: GRANULINS HAVE POSSIBLE CYTOKINE-LIKE ACTIVITY. THEY MAY



FT CARBOHYD 368 368 N-LINKED (GLCNAC. . .)  
FT CARBOHYD 530 530 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT VARSPLIC 377 531 MISSING (IN ISOFORM 2).  
FT VARIANT 454 454 G -> O.  
FT /FTID-VAR\_003445.  
FT CONFLICT 219 219 S -> H (IN REF. 6).  
FT CONFLICT 386 386 W -> H (IN REF. 6).  
SQ SEQUENCE 593 AA; 63473 MW; 4E402BDB16DE2819 CRC64;  
  
Query Match 73.9%; Score 51; DB 1; Length 593;  
Best Local Similarity 56.7%; Pred. No. 0.38;  
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 PDKATQCPDST 12  
IIII::IIII II  
DB 209 PDARSRCPDGST 220  
  
RESULT 5  
CDC7\_MOUSE STANDARD; PRT; 564 AA.  
ID CDC7\_MOUSE Q922Y7;  
AC Q920H0: Q922Y7;  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DE 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Cell division cycle 7-related protein kinase (EC 2.7.1.1-) (CDC7-  
related kinase) (muCdc7).  
GN CDC7L1 OR CDC7.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=98389757; PubMed=9722556;  
RA Kim J., Sato N., Yamada M., Arai K., Masai H.;  
RT "Growth regulation of the expression of mouse cDNA and gene encoding a  
serine/threonine kinase related to Saccharomyces cerevisiae CDC7  
essential for G1/S transition. Structure, chromosomal localization,  
and expression of mouse gene for S. cerevisiae Cdc7-related kinase.,"  
J. Biol. Chem. 273:23248-23257(1998).  
CC THE G1/S PHASE TRANSITION AND/OR DNA REPLICATION. CAN  
CC PHOSPHORYLATES MC2 AND MC3.  
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).  
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS; 1 (SHOWN HERE) AND 2;  
CC ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CC CDC7 SUBFAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; AB019388; BAA34347.1;  
CC EMBL; AB019378; BAA34347.1; JOINED.  
CC EMBL; AB019379; BAA34347.1; JOINED.  
CC EMBL; AB019380; BAA34347.1; JOINED.  
CC EMBL; AB019381; BAA34347.1; JOINED.  
CC EMBL; AB019382; BAA34347.1; JOINED.  
CC EMBL; AB019383; BAA34347.1; JOINED.  
CC EMBL; AB019384; BAA34347.1; JOINED.  
CC EMBL; AB019385; BAA34347.1; JOINED.  
CC EMBL; AB019386; BAA34347.1; JOINED.  
CC EMBL; AB019387; BAA34347.1; JOINED.  
CC EMBL; AB018575; BAA33881.1;  
CC EMBL; AB018574; BAA33880.1;  
CC MGD; MGI:1309511; Cdc7l1.

DR InterPro: IPR000719; Euk\_pkinase.  
DR InterPro: IPR002290; Ser\_thr\_pkinase.  
DR Pfam: PF00069; pkinase; 1.  
DR ProDom: PD000001; Euk\_pkinase; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
KW Cell cycle; Cell division; Serine/threonine-protein kinase;  
KW Transferase; ATP-binding; Alternative splicing; Nuclear protein.  
FT DOMAIN 52 564 PROTEIN KINASE.  
FT NP\_BIND 58 66 ATP (BY SIMILARITY).  
FT BINDING 84 84 ATP (BY SIMILARITY).  
FT ACT\_SITE 171 171 BY SIMILARITY.  
FT VARSPLIC 267 298 MISSING (IN ISOFORM 2).  
SQ SEQUENCE 564 AA; 62752 MW; 925791D5F93FBA02 CRC64;  
  
Query Match 59.4%; Score 41; DB 1; Length 564;  
Best Local Similarity 50.0%; Pred. No. 16;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 PDKATQCPDST 12  
II::IIII II  
DB 373 PEVLTKCPDQTT 384  
  
RESULT 6  
FRUA\_STRMO STANDARD; PRT; 1423 AA.  
ID FRUA\_STRMO Q03174;  
AC Q03174;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Fructan beta-fructosidase precursor (EC 3.2.1.80) (Exo-beta-D-  
fructosidase) (fructanase).  
GN FRUA.  
OS Streptococcus mutans.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1309;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=GS-5;  
RX MEDLINE=93014178; PubMed=1398976;  
RA Burne R.A., Penders J.E.C.;  
RT "Characterization of the Streptococcus mutans GS-5 frua gene encoding  
exo-beta-D-fructosidase".  
RL Infect. Immun. 60:4621-4632(1992).  
CC -1- FUNCTION: THIS PROTEIN IS A FRUCTANASE ENZYME WHICH DEGRADES  
CC LEVANS AND INULINS TO FRUCTOSE AND ALSO CLEAVES SUCROSE INTO  
CC GLUCOSE AND FRUCTOSE AND CAN THEREFORE FUNCTION AS AN  
CC EXTRACELLULAR INVERTASE.  
CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal-non-reducing 2,1- and  
CC 2,6-linked beta-D-fructofuranose residues in fructans.  
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan  
CC by an amide bond (Potential).  
CC -1- INDUCTION: BY SUCROSE, FRUCTAN SUBSTRATES AND FRUCTOSE.  
CC -1- SIMILARITY: BELONGS TO FAMILY 32 OF GLYCOSYL HYDROLASES.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; U78296; AAA26889.1;  
CC InterPro: IPR003343; Big\_2.  
CC InterPro: IPR001362; GH\_32.  
CC InterPro: IPR001899; Gram\_pos\_anchor.  
CC Pfam; PF00251; Glyco\_hydro\_32; 1.  
CC Pfam; PF02368; Big\_2; 1.  
CC TIGRFAMS; TIGR01167; LPXTG\_anchor; 1.

DR PROSITE; PS00609; GLYCOSYL\_HYDROL\_F32; FALSE\_NEG.  
 DR PROSITE; PS00847; GRAM\_POS\_ANCHORING; FALSE\_NEG.  
 KW Hydrolase; Glycosidase; Cell wall; Peptidoglycan-anchor; Signal.  
 FT SIGNAL 1 39  
 FT CHAIN 40 1391  
 FT PROPEP 1392 1423  
 FT ACT\_SITE 458 458  
 FT DOMAIN 867 871  
 FT SITE 1388 1392  
 FT MOD\_RES 1391 1391  
 SQ SEQUENCE 1423 AA; 158668 MW; 8E574715F4E72A8A CRC64;  
 Query Match 58.0%; Score 40; DB 1; Length 1423;  
 Best Local Similarity 66.7%; Pred. No. 56;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 PDAKTCPCD 9  
 DB 619 PDLHTECPD 627  
 RESULT 7  
 RNTR\_TRIVI  
 ID RNTR\_TRIVI STANDARD; PRT; 234 AA.  
 AC P24657;  
 DT 01-MAR-1992 (Rel. 21, Created)  
 DT 01-MAR-1992 (Rel. 21, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Ribonuclease Trv (EC 3.1.27.1) (RNase Trv).  
 OS Trichoderma viride.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Hypocreales; mitosporic Hypocreales; Trichoderma.  
 OX NCBI\_TaxID=5547;  
 RN SEQUENCE.  
 RP MEDLINE=92176163; PubMed=1794979;  
 RA Inada Y., Watanabe H., Ohgi K., Irie M.;  
 RT "Isolation, characterization, and primary structure of a base non-specific and adenylic acid preferential ribonuclease with higher specific activity from Trichoderma viride.";  
 RL J. Biochem. 110:896-904(1991).  
 CC -1- FUNCTION: THIS IS A BASE NON-SPECIFIC AND ADENYLIC ACID PREFERENTIAL RIBONUCLEASE.  
 CC -1- CATALYTIC ACTIVITY: Two-stage endonucleolytic cleavage to nucleoside 3'-phosphates and 3'-phosphooligonucleotides with 2',3'-cyclic phosphate intermediates.  
 CC -1- SIMILARITY: BELONGS TO THE RNASE T2 FAMILY.  
 CC PIR; JK0197; JK0197.  
 DR HSP; P08056; IBOI.  
 DR InterPro: IPR001568; RNase\_T2.  
 DR Pfam; PF00445; ribonuclease\_T2; 1.  
 DR PROSITE; PS00530; RNase\_T2\_1; 1.  
 DR PROSITE; PS00531; RNase\_T2\_2; 1.  
 KW Hydrolase; Nuclease; Endonuclease; Glycoprotein.  
 FT DISULFID 5 24  
 FT DISULFID 13 59  
 FT DISULFID 23 125  
 FT DISULFID 67 117  
 FT DISULFID 189 224  
 FT ACT\_SITE 52 52  
 FT ACT\_SITE 110 110  
 FT ACT\_SITE 114 114  
 FT CARBOHYD 15 15  
 FT CARBOHYD 75 75  
 SQ SEQUENCE 234 AA; 25901 MW; 8B011DD65A3F909A CRC64;  
 Query Match 57.2%; Score 39.5; DB 1; Length 234;  
 Best Local Similarity 72.7%; Pred. No. 12;  
 Matches 8; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 PD-AKTQCPDD 10  
 DB 217 PDGAKSTCPDD 227  
 RESULT 8  
 GRNL\_CYPCA  
 ID GRNL\_CYPCA STANDARD; PRT; 57 AA.  
 AC P81013;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE Granulin 1.  
 OS Cyprinus carpio (Common carp).  
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Cyprinus.  
 OX NCBI\_TaxID=7962;  
 RN SEQUENCE.  
 RP TISSUE=Spleen, and Kidney;  
 RX MEDLINE=93252781; PubMed=8486624;  
 RA Belcourt D.R., Lazure C., Bennett R.P.;  
 RT "Isolation and primary structure of the three major forms of granulin-like peptides from hematopoietic tissues of a teleost fish (Cyprinus carpio).";  
 RL J. Biol. Chem. 268:9230-9237(1993).  
 CC -1- FUNCTION: GRANULINS HAVE POSSIBLE CYTOKINE-LIKE ACTIVITY. THEY MAY PLAY A ROLE IN INFLAMMATION, WOUND REPAIR, AND TISSUE REMODELING.  
 CC -1- TISSUE SPECIFICITY: UBIQUITOUS.  
 CC -1- PTM: GRANULINS ARE DISULFIDE BRIDGED.  
 DR InterPro: IPR000118; Granulin.  
 DR Pfam; PF00396; granulin; 1.  
 DR SMART; SM00277; GRAN; 1.  
 KW Cytokine.  
 SQ SEQUENCE 57 AA; 6289 MW; E4A131B1288FE55A CRC64;  
 Query Match 55.1%; Score 38; DB 1; Length 57;  
 Best Local Similarity 63.6%; Pred. No. 5.6;  
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 DAKTQCPDDST 12  
 DB 5 DAATTCPDGTT 15  
 RESULT 9  
 VS12\_TRYBB  
 ID VS12\_TRYBB STANDARD; PRT; 479 AA.  
 AC P26327;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Variant surface glycoprotein ILTAT 1.22 precursor (VSG).  
 OS Trypanosoma brucei brucei.  
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.  
 OX NCBI\_TaxID=5702;  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=Isolate MIAG 202B;  
 RX MEDLINE=92046037; PubMed=1942032;  
 RA Carrington M., Miller N., Blum M.L., Roditi I., Wiley D.C., Turner M.J.;  
 RT "Variant specific glycoprotein of Trypanosoma brucei consists of two domains each having an independently conserved pattern of cysteine residues.";  
 RL J. Mol. Biol. 221:823-835(1991).  
 CC -1- FUNCTION: VSG FORMS A COAT ON THE SURFACE OF THE PARASITE. THE TYMANOSOME EVADES THE IMMUNE RESPONSE OF THE HOST BY EXPRESSING A SERIES OF ANTIGENICALLY DISTINCT VSGS FROM AN ESTIMATED 1000 VSG GENES.  
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor. A SOLUBLE FORM IS RELEASED FROM RUPTURED CELLS BY THE ACTION OF A



DE GN MYOblast determination protein 1.  
 OS MYO1 OR MYO.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88080485; PubMed=3690668;  
 RA Davis R.L., Weintraub H., Lassar A.B.;  
 RT "Expression of a single transcribed cDNA converts fibroblasts to  
 myoblasts.";  
 RL Cell 51:987-1000(1987).  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92093599; PubMed=1754380;  
 RA Zingg J.M., Alva G.P., Jost J.P.;  
 RT "Characterisation of a genomic clone covering the structural mouse  
 MyoD1 gene and its promoter region.";  
 RL Nucleic Acids Res. 19:6433-6439(1991).  
 RN (3)  
 RP FUNCTION.  
 RX MEDLINE=89019371; PubMed=3175662;  
 RA Tapscott S.J., Davis R.L., Thayer M.J., Cheng P.-F., Weintraub H.,  
 RA Lassar A.B.;  
 RT "MyoD1: a nuclear phosphoprotein requiring a Myc homology region to  
 convert fibroblasts to myoblasts.";  
 RL Science 242:405-411(1988).  
 RN (4)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88223371; PubMed=3286015;  
 RA Pinney D.F., Pearson-White S.H., Konieczny S.F., Latham K.E.,  
 RA Emerson C.P., Jr.;  
 RT "Myogenic lineage determination and differentiation: evidence for a  
 regulatory gene pathway.";  
 RL Cell 53:781-793(1988).  
 RN (5)  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 102-166.  
 RX MEDLINE=94236689; PubMed=8181063;  
 RA Ma P.C.M., Rould M.A., Weintraub H., Pabo C.O.;  
 RT "Crystal structure of MyoD bHLH domain-DNA complex: perspectives on  
 DNA recognition and implications for transcriptional activation.";  
 RL Cell 77:451-459(1994).  
 RN (6)  
 RP INHIBITION BY TWIST.  
 RX MEDLINE=98001585; PubMed=9343420;  
 RA Hamamori Y., Wu H.Y., Sartorelli V., Kedes L.;  
 RT "The basic domain of myogenic basic helix-loop-helix (bHLH) proteins  
 is the novel target for direct inhibition by another bHLH protein,  
 Twist.";  
 RL Mol. Cell. Biol. 17:6563-6573(1997).  
 CC -1- FUNCTION: INVOLVED IN MUSCLE DIFFERENTIATION (MYOGENIC FACTOR).  
 CC INDUCES FIBROBLASTS TO DIFFERENTIATE INTO MYOBLASTS. ACTIVATES  
 CC MUSCLE-SPECIFIC PROMOTERS. INTERACTS WITH AND IS INHIBITED BY THE  
 CC TWIST PROTEIN. THIS INTERACTION PROBABLY INVOLVES THE BASIC  
 CC DOMAINS OF BOTH PROTEINS.  
 CC -1- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER  
 CC bHLH PROTEIN. SEEMS TO FORM ACTIVE HETERODIMERS WITH ITP-2.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (bHLH) FAMILY OF  
 CC TRANSCRIPTION FACTORS. "MYOGENIC FACTORS" SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; M18779; AAA39799.1; -  
 DR EMBL; X61655; CAA43836.1; -  
 DR EMBL; M84918; AAA39798.1; -

DR PIR; A29636; A29636.  
 DR PIR; S22587; S22587.  
 DR PDB; 1MDY; 31-AUG-94.  
 DR TRANSFAC; T00526; -  
 DR MGD; MGI:97275; MyoD1.  
 DR InterPro; IPR002546; Basic.  
 DR InterPro; IPR001092; HLH\_basic.  
 DR Pfam; PF00010; HLH; 1.  
 DR Pfam; PF01586; Basic; 1.  
 DR SMART; SM00520; BASIC; 1.  
 DR SMART; SM00353; HLH; 1.  
 DR PROSITE; PS00038; HLH\_1; 1.  
 DR PROSITE; PS50888; HLH\_2; 1.  
 KW Myogenesis; Differentiation; Developmental protein; Nuclear protein;  
 KW Transcription regulation; DNA-binding; 3D-structure.  
 FT DNA\_BIND 109 121 BASIC DOMAIN  
 FT DOMAIN 122 161 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).  
 SQ SEQUENCE 318 AA; 48C836896FCBDC0C CRC64;  
 Query Match 53.6%; Score 37; DB 1; Length 318;  
 Best Local Similarity 63.6%; Pred. No. 42;  
 Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 PDAKTQCDDDS 11  
 DB 299 PDAAPQCPAGS 309  
 RESULT 13  
 N214\_HUMAN  
 ID N214\_HUMAN STANDARD; PRT; 2090 AA.  
 AC P35658;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Nuclear pore complex protein Nup214 (Nucleoporin Nup214) (214 kDa  
 DE nucleoporin) (CAN protein).  
 GN NUP214 OR CAN OR CAIN.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RX MEDLINE=92195315; PubMed=1549122;  
 RA Von Lindern M., Fornerod M., Van Baal S., Jaegle M., De Wit T.,  
 RA Buljys A., Grosveld G.;  
 RT "The translocation (6;9), associated with a specific subtype of acute  
 RT myeloid leukemia, results in the fusion of two genes, dek and can,  
 RT and the expression of a chimeric, leukemia-specific dek-can mRNA.";  
 RL Mol. Cell. Biol. 12:1687-1697(1992).  
 RN [2]  
 RP CHARACTERIZATION.  
 RX MEDLINE=94151361; PubMed=8108440;  
 RA Krammer D., Wozniak R.W., Blobel G., Radu A.;  
 RT "The human CAN protein, a putative oncogene product associated with  
 RT myeloid leukemogenesis, is a nuclear pore complex protein that faces  
 RT the cytoplasm.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:1519-1523(1994).  
 CC -1- FUNCTION: MAY SERVE AS A DOCKING SITE IN THE RECEPTOR-MEDIATED  
 CC IMPORT OF SUBSTRATES ACROSS THE NUCLEAR PORE COMPLEX.  
 CC -1- SUBUNIT: HOMODIMER. INTERACTS WITH DDX19.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR PORE COMPLEX. CYTOPLASMIC FILAMENTS.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THYMUS, SPLEEN, BONE MARROW,  
 CC KIDNEY, BRAIN AND TESTIS, BUT HARDLY IN ALL OTHER TISSUES OR IN  
 CC WHOLE EMBRYOS DURING DEVELOPMENT.  
 CC -1- DOMAIN: CONTAINS F-G REPEATS.  
 CC -1- PTM: PROBABLY GLYCOSYLATED AS IT REACTS WITH WHEAT GERM AGGLUTININ  
 CC (WGA).  
 CC -1- DISEASE: IMPLICATED IN A SUBSET OF ACUTE MYELOID LEUKEMIA (ACUTE  
 CC NONLYMPHOBLASTIC LEUKEMIA) (AML) CARRYING A CHROMOSOMAL  
 CC TRANSLOCATION T(6;9)(P23;Q34) THAT RESULTS IN THE FORMATION OF A

CC DEK-CAN FUSION GENE.  
 CC -1- DISEASE: IN A CASE OF ACUTE UNDIFFERENTIATED LEUKEMIA (AUL) A  
 CC TRANSLOCATION RESULTS IN THE FORMATION OF A SET-CAN FUSION GENE.  
 CC -1- DATABASE: NAME-Atlas Genet. Cytogetnet. Oncol. Haematol.;  
 CC WWW="http://www.infobiogen.fr/services/chronocancer/Genes/CAN.html".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; X64228; CAA45535.1; -  
 CC PIR; S26058; S26058.  
 CC Genew; HGNC:8064; NUP214.  
 CC MIM; 114350; -  
 CC InterPro; IPR004325; Nucleoporin\_FG.  
 CC InterPro; IPR001680; WD40.  
 CC Pfam; PF03093; Nucleoporin\_FG; 17.  
 CC SMART; SM00320; WD40; 1.  
 CC Nuclear protein; Transport; Proto-oncogene; Chromosomal translocation;  
 CC Repeat; Glycoprotein.  
 CC DOMAIN 481 2076 11 X 5 AA APPROXIMATE REPEATS.  
 CC FT DOMAIN 1409 2084 18 X 4 AA APPROXIMATE REPEATS.  
 CC FT DOMAIN 1427 2085 11 X 3 AA APPROXIMATE REPEATS.  
 CC FT DOMAIN 1213 2090 PRO/SER/THR-RICH.  
 CC FT DOMAIN 740 768 LEUCINE-ZIPPER 1.  
 CC FT DOMAIN 861 882 LEUCINE-ZIPPER 2.  
 CC FT SITE 812 813 BREAKPOINT.  
 CC SQ SEQUENCE 2090 AA; 213766 MW; 6DBE767FDD857F8F CRC64;  
 CC -----  
 CC Query Match 53.6%; Score 37; DB 1; Length 2090;  
 CC Best Local Similarity 66.7%; Pred. No. 2.5e+02;  
 CC Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 CC -----  
 CC QY 2 DARTQCPDD 10  
 CC II : I I I  
 CC 998 DARTSCKDD 1006  
 CC -----  
 CC RESULT 14  
 CC AHNK\_HUMAN STANDARD; PRT; 2960 AA.  
 CC AC Q09666;  
 CC DT 01-FEB-1996 (Rel. 33, Created)  
 CC DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 CC DE Neuroblast differentiation associated protein AHNAK (Desmoyokin)  
 CC DE (Fragments).  
 CC GN AHNAK OR PM227.  
 CC OS Homo sapiens (Human).  
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 CC OX NCBI\_TaxID=9606;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RC TISSUE=Placenta;  
 CC RX MEDLINE=92302265; PubMed=1608957;  
 CC RA Shivelman E., Cohen F.E., Bishop J.M.;  
 CC RT "A human gene (AHNAK) encoding an unusually large protein with a 1.2-  
 CC RT microns polyanionic rod structure.";  
 CC RL Proc. Natl. Acad. Sci. U.S.A. 89:5472-5476(1992).  
 CC [2]  
 CC RP ERRATUM  
 CC RA Shivelman E., Cohen F.E., Bishop J.M.;  
 CC RL Proc. Natl. Acad. Sci. U.S.A. 90:4328-4328(1993).  
 CC -1- FUNCTION: MAY BE REQUIRED FOR NEURONAL CELL DIFFERENTIATION.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-62 IS THE  
 CC INITIATOR.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; M80902; AAA69899.1; -  
 CC EMBL; M80899; AAA69898.1; -  
 CC Genew; HGNC:347; AHNAK.  
 CC MIM; 103390; -  
 CC Nuclear protein; Repeat.  
 CC KW NON\_CONS 1583 1684  
 CC FT DOMAIN 2041 2049 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 CC FT DOMAIN 2089 2097 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 CC FT DOMAIN 2104 2109 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 CC FT DOMAIN 2776 2786 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 CC FT DOMAIN 2842 2849 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 CC SQ SEQUENCE 2960 AA; 312487 MW; F9FDD2103A35232A CRC64;  
 CC -----  
 CC Query Match 53.6%; Score 37; DB 1; Length 2960;  
 CC Best Local Similarity 55.8%; Pred. No. 3.4e+02;  
 CC Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 CC -----  
 CC QY 1 PDAKTCPD 9  
 CC II : I I I  
 CC 1647 PDVELECPD 1655  
 CC -----  
 CC RESULT 15  
 CC TENX\_HUMAN STANDARD; PRT; 4289 AA.  
 CC AC P22105; P78531; Q08424; Q9UMG7;  
 CC DT 01-AUG-1991 (Rel. 19, Created)  
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 CC DE Tenascin-X precursor (TN-X) (Hexabrachion-like).  
 CC GN TNXB OR TNX OR XB OR HXBL.  
 CC OS Homo sapiens (Human).  
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 CC OX NCBI\_TaxID=9606;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RA Rowen L., Dankers C., Baskin D., Faust J., Loretz C., Ahearn M.E.,  
 CC RA Banta A., Schwartzell S., Smith T.M., Spies T., Hood L.;  
 CC RT "Sequence determination of 300 kilobases of the human class III MHC  
 CC RT locus.";  
 CC RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 CC [2]  
 CC RP SEQUENCE OF 1-747 AND 1687-1944 FROM N.A.  
 CC RC TISSUE=Leukocyte;  
 CC RX MEDLINE=93300909; PubMed=7686164;  
 CC RA Bristow J., Tee M.K., Gitelman S.E., Mellon S.H., Miller W.L.;  
 CC RT "Tenascin-X: a novel extracellular matrix protein encoded by the human  
 CC RT XB gene overlapping P450c21B.";  
 CC RL J. Cell Biol. 122:265-278(1993).  
 CC [3]  
 CC RP SEQUENCE FROM N.A. (ISOFORM XB-SHORT).  
 CC RC TISSUE=Adrenal gland;  
 CC RX MEDLINE=96015044; PubMed=8530023;  
 CC RA Tee M.K., Thomson A.A., Bristow J., Miller W.L.;  
 CC RT "Sequences promoting the transcription of the human XA gene  
 CC RT overlapping P450c21A correctly predict the presence of a novel,  
 CC RT adrenal-specific, truncated form of tenascin-X.";  
 CC RL Genomics 28:171-178(1995).  
 CC [4]  
 CC RP SEQUENCE OF 1-23 FROM N.A.  
 CC RC TISSUE=Fetal adrenal gland;  
 CC RX MEDLINE=97081760; PubMed=8923003;  
 CC RA Speck M., Barry F., Miller W.L.;  
 CC RT "Alternate promoters and alternate splicing of human tenascin-X, a

gene with 5' and 3' ends buried in other genes.";

[5] Hum. Mol. Genet. 5:1749-1758(1996).

RP SEQUENCE OF 3470-4289 FROM N.A.

RA MEDLINE=89367293; PubMed=2475872;

RA Morel Y., Bristow J., Gitelman S.E., Miller W.L.;

RT "transcript encoded on the opposite strand of the human steroid 21-

RT hydroxylase/complement component C4 gene locus.";

RA Proc. Natl. Acad. Sci. U.S.A. 86:6582-6586(1989).

RP [6]

RP DISEASE.

RA PubMed=11642233;

RA Schalkwijk J., Zweers M.C., Steijlen P.M., Dean W.B., Taylor G.,

RA van Vlijmen I.M., van Haren B., Miller W.L., Bristow J.;

RT "A recessive form of the Ehlers-Danlos syndrome caused by tenascin-X

RT deficiency";

RL New Engl. J. Med. 345:1167-1175(2001).

CC -1- FUNCTION: APPEARS TO MEDIATE INTERACTIONS BETWEEN CELLS AND THE

CC EXTRACELLULAR MATRIX. SUBSTRATE-ADHESION MOLECULE THAT APPEARS TO

CC INHIBIT CELL MIGRATION. MAY PLAY A ROLE IN SUPPORTING THE GROWTH

CC OF EPITHELIAL TUMORS.

CC -1- SUBCELLULAR LOCATION: Extracellular matrix.

CC -1- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS; XB (SHOWN HERE) AND

CC XB-SHORT; ARE PRODUCED BY ALTERNATIVE SPLICING.

CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN FETAL ADRENAL, IN FETAL

CC TESTIS, FETAL SMOOTH, STRIATED AND CARDIAC MUSCLE. XB-SHORT IS

CC ONLY EXPRESSED IN THE ADRENAL GLAND.

CC -1- DISEASE: ASSOCIATION WITH CONGENITAL ADRENAL HYPERPLASIA.

CC -1- DISEASE: Defects in TNXB are the cause of an autosomal recessive

CC form of Ehlers-Danlos syndrome (EDS) manifesting with

CC hyperextensible skin, hypermobile joints, and tissue fragility as

CC observed in classical EDS, but lacking atrophic scars and delayed

CC wound healing.

CC -1- SIMILARITY: CONTAINS 19 EGF-LIKE DOMAINS.

CC -1- SIMILARITY: CONTAINS 32 FIBRONECTIN TYPE III-LIKE DOMAINS.

CC -1- SIMILARITY: CONTAINS 1 FIBRINOGEN C-TERMINAL DOMAIN.

CC -1- CAUTION: THERE ARE TWO GENES FOR TN-X: TNXA AND TNXB. TNXA IS A

CC PARTIAL GENE WHICH CAN SOMETIMES RECOMBINE WITH TNXB.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; U89337; AAB47488.1; -

DR EMBL; AF019413; AAB67981.1; -

DR EMBL; X71923; CAA50739.1; -

DR EMBL; Y13782; CAA74109.1; -

DR EMBL; Y13783; CAA74110.1; -

DR EMBL; U24488; AAB41287.1; -

DR EMBL; U52696; AAC50889.1; -

DR EMBL; M25813; AAA35884.1; -

DR HSSP; P02671; 1FZD.

DR Genew; HGNC:11976; TNXB.

DR MIN; 600985; -

DR MIN; 606408; -

DR InterPro; IPR000561; EGF-like.

DR InterPro; IPR003961; FN\_III.

DR InterPro; IPR002181; Fibrinogen\_C.

DR InterPro; IPR002049; Laminin\_EGF.

DR Pfam; PF00008; EGF; 15.

DR Pfam; PF00041; fn3; 33.

DR Pfam; PF00147; fibrinogen\_C; 1.

DR PRINTS; PR00011; EGF\_LAMININ.

DR SMART; SM00181; EGF; 8.

DR SMART; SM00001; EGF\_like; 1.

DR SMART; SM00186; FBG; 1.

DR SMART; SM00060; FN3; 31.

DR PROSITE; PS00022; EGF\_1; 18.

DR PROSITE; PS01186; EGF\_2; 19.

DR PROSITE; PS00514; FIBRIN\_AG\_C\_DOMAIN; 1.

KW Glycoprotein; Cell adhesion; Repeat; EGF-like domain; Coiled coil;

KW Extracellular matrix; Alternative splicing; Signal;

KW Ehlers-Danlos syndrome.

FT SIGNAL 1 23 POTENTIAL.

FT CHAIN 24 4289 TENASCIN-X.

FT DOMAIN 156 168 EGF-LIKE 1 (INCOMPLETE).

FT DOMAIN 183 213 EGF-LIKE 2.

FT DOMAIN 214 244 EGF-LIKE 3.

FT DOMAIN 245 275 EGF-LIKE 4.

FT DOMAIN 276 306 EGF-LIKE 5.

FT DOMAIN 307 337 EGF-LIKE 6.

FT DOMAIN 338 368 EGF-LIKE 7.

FT DOMAIN 369 399 EGF-LIKE 8.

FT DOMAIN 400 430 EGF-LIKE 9.

FT DOMAIN 431 461 EGF-LIKE 10.

FT DOMAIN 462 492 EGF-LIKE 11.

FT DOMAIN 493 523 EGF-LIKE 12.

FT DOMAIN 524 554 EGF-LIKE 13.

FT DOMAIN 555 585 EGF-LIKE 14.

FT DOMAIN 586 616 EGF-LIKE 15.

FT DOMAIN 617 647 EGF-LIKE 16.

FT DOMAIN 648 679 EGF-LIKE 17.

FT DOMAIN 684 714 EGF-LIKE 18.

FT DOMAIN 715 746 EGF-LIKE 19.

FT DOMAIN 792 872 FIBRONECTIN TYPE-III 1.

FT DOMAIN 901 922 COILED COIL (POTENTIAL).

FT DOMAIN 941 1021 FIBRONECTIN TYPE-III 2.

FT DOMAIN 1047 1127 FIBRONECTIN TYPE-III 3.

FT DOMAIN 1149 1226 FIBRONECTIN TYPE-III 4.

FT DOMAIN 1246 1327 FIBRONECTIN TYPE-III 5.

FT DOMAIN 1348 1429 FIBRONECTIN TYPE-III 6.

FT DOMAIN 1459 1540 FIBRONECTIN TYPE-III 7.

FT DOMAIN 1561 1642 FIBRONECTIN TYPE-III 8.

FT DOMAIN 1659 1736 FIBRONECTIN TYPE-III 9.

FT DOMAIN 1756 1836 FIBRONECTIN TYPE-III 10.

FT DOMAIN 1856 1939 FIBRONECTIN TYPE-III 11.

FT DOMAIN 1962 2039 FIBRONECTIN TYPE-III 12.

FT DOMAIN 2069 2150 FIBRONECTIN TYPE-III 13.

FT DOMAIN 2167 2248 FIBRONECTIN TYPE-III 14.

FT DOMAIN 2266 2347 FIBRONECTIN TYPE-III 15.

FT DOMAIN 2365 2446 FIBRONECTIN TYPE-III 16.

FT DOMAIN 2471 2552 FIBRONECTIN TYPE-III 17.

FT DOMAIN 2562 2663 FIBRONECTIN TYPE-III 18.

FT DOMAIN 2688 2769 FIBRONECTIN TYPE-III 19.

FT DOMAIN 2794 2875 FIBRONECTIN TYPE-III 20.

FT DOMAIN 2889 2972 FIBRONECTIN TYPE-III 21.

FT DOMAIN 2997 3078 FIBRONECTIN TYPE-III 22.

FT DOMAIN 3105 3186 FIBRONECTIN TYPE-III 23.

FT DOMAIN 3211 3292 FIBRONECTIN TYPE-III 24.

FT DOMAIN 3307 3384 FIBRONECTIN TYPE-III 25.

FT DOMAIN 3399 3481 FIBRONECTIN TYPE-III 26.

FT DOMAIN 3494 3575 FIBRONECTIN TYPE-III 27.

FT DOMAIN 3601 3682 FIBRONECTIN TYPE-III 28.

FT DOMAIN 3699 3787 FIBRONECTIN TYPE-III 29.

FT DOMAIN 3801 3879 FIBRONECTIN TYPE-III 30.

FT DOMAIN 3890 3971 FIBRONECTIN TYPE-III 31.

FT DOMAIN 3978 4059 FIBRONECTIN TYPE-III 32.

FT DOMAIN 4071 4289 FIBRINOGEN C-TERMINAL.

FT SITE 1748 1750 CELL ATTACHMENT SITE (POTENTIAL).

FT DISULFID 187 197 BY SIMILARITY.

FT DISULFID 191 202 BY SIMILARITY.

FT DISULFID 204 213 BY SIMILARITY.

FT DISULFID 218 228 BY SIMILARITY.

FT DISULFID 222 233 BY SIMILARITY.

FT DISULFID 235 244 BY SIMILARITY.

FT DISULFID 249 259 BY SIMILARITY.

FT DISULFID 253 264 BY SIMILARITY.

FT DISULFID 266 275 BY SIMILARITY.

FT DISULFID 280 290 BY SIMILARITY.

FT DISULFID 284 295 BY SIMILARITY.

FT DISULFID 297 306 BY SIMILARITY.

FT DISULFID 311 321 BY SIMILARITY.



FT	DISULFID	315	326	BY SIMILARITY.
FT	DISULFID	328	337	BY SIMILARITY.
FT	DISULFID	342	352	BY SIMILARITY.
FT	DISULFID	346	357	BY SIMILARITY.
FT	DISULFID	359	368	BY SIMILARITY.
FT	DISULFID	373	383	BY SIMILARITY.
FT	DISULFID	377	388	BY SIMILARITY.
FT	DISULFID	390	399	BY SIMILARITY.
FT	DISULFID	404	414	BY SIMILARITY.
FT	DISULFID	408	419	BY SIMILARITY.
FT	DISULFID	421	430	BY SIMILARITY.
FT	DISULFID	435	445	BY SIMILARITY.
FT	DISULFID	439	450	BY SIMILARITY.

Query Match 53.6%; Score 37; DB 1; Length 4289;  
Best Local Similarity 60.0%; Pred No. 4.9e+02;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 PDAKTQCPDD 10  
Db 181 PSAGSGCPDD 190

Search completed: July 8, 2003, 16:25:01  
Job time : 7.84615 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2003, 16:17:10 ; Search time 25.8462 Seconds  
(without alignments)  
95.665 Million cell updates/sec

Title: US-09-824-647-4

Perfect score: 69

Sequence: 1 PDARTQCPDDST 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	69	100.0	589	11 Q9D2V3	Q9D2V3 mus musculus
2	65	94.2	51	11 Q9QWB4	Q9QWB4 mus sp. pc
3	47	68.1	384	5 Q8T9J3	Q8T9J3 drosophila
4	45	65.2	511	12 Q89001	Q89001 vesicular s
5	45	65.2	511	12 Q89002	Q89002 vesicular s
6	45	65.2	511	12 Q89005	Q89005 vesicular s
7	45	65.2	511	12 Q89008	Q89008 vesicular s
8	45	65.2	511	12 Q89625	Q89625 vesicular s
9	45	65.2	511	12 Q89767	Q89767 vesicular s
10	45	65.2	511	12 Q89878	Q89878 vesicular s
11	45	65.2	1557	5 Q8WPK9	Q8WPK9 olkopleura
12	43	62.3	358	5 Q8U362	Q8U362 caenorhabdi
13	42	60.9	463	17 Q8ZTL3	Q8ZTL3 pyrobaculum
14	42	60.9	506	16 Q8YD93	Q8YD93 brucella me
15	41	59.4	147	13 Q90ZD0	Q90ZD0 brachydanio
16	41	59.4	147	13 Q90ZC9	Q90ZC9 brachydanio

17	41	59.4	147	13	Q90ZC8	Q90ZC8 brachydanio
18	41	59.4	564	11	Q9W0V1	Q9W0V1 mus musculus
19	40	58.0	353	12	Q9PYZ0	Q9PYZ0 xestia c-ni
20	39	56.5	190	5	Q76599	Q76599 caenorhabdi
21	39	56.5	222	5	Q20511	Q20511 caenorhabdi
22	39	56.5	465	16	Q92JG3	Q92JG3 rickettsia
23	38	55.1	59	4	Q9PIL5	Q9PIL5 homo sapien
24	38	55.1	144	17	Q9HRV3	Q9HRV3 halobacteri
25	38	55.1	351	16	Q92FS3	Q92FS3 listeria in
26	38	55.1	351	16	Q8YAT2	Q8YAT2 listeria in
27	38	55.1	436	17	Q29289	Q29289 archaeoglob
28	38	55.1	479	5	Q96766	Q96766 trypanosoma
29	38	55.1	480	5	Q76979	Q76979 trypanosoma
30	38	55.1	483	13	Q13045	Q13045 xenopus lae
31	38	55.1	505	3	Q9P8P3	Q9P8P3 trichoderma
32	38	55.1	511	12	Q88996	Q88996 vesicular s
33	38	55.1	511	12	Q88997	Q88997 vesicular s
34	38	55.1	511	12	Q88998	Q88998 vesicular s
35	38	55.1	511	12	Q89000	Q89000 vesicular s
36	38	55.1	511	12	Q89003	Q89003 vesicular s
37	38	55.1	511	12	Q89004	Q89004 vesicular s
38	38	55.1	511	12	Q89007	Q89007 vesicular s
39	38	55.1	511	12	Q89664	Q89664 vesicular s
40	38	55.1	511	12	Q89850	Q89850 vesicular s
41	38	55.1	706	16	Q9A2D2	Q9A2D2 caulobacter
42	38	55.1	993	12	Q90Z27	Q90Z27 cucumber mo
43	38	55.1	1125	5	P92135	P92135 entamoeba d
44	38	55.1	1283	5	Q18824	Q18824 caenorhabdi
45	37	53.6	127	16	Q8UGM0	Q8UGM0 agrobacteri

#### ALIGNMENTS

#### RESULT 1

Q9D2V3	ID	Q9D2V3	PRELIMINARY	PRT	589 AA.
AC	Q9D2V3				
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
DT	01-JUN-2001	(TREMBLrel. 21, Last annotation update)			
DE	Adult male kidney cDNA, RIKEN full-length enriched library,				
DE	clone:0610012H06, full insert sequence.				
GN	GRN.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=C57BL/6J; TISSUE=KIDNEY;				
RX	MEDLINE=21085660; PubMed=11217851;				
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,				
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,				
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,				
RA	Saito T., Okazaki Y., Gofobori T., Bono H., Kasukawa T., Saito R.,				
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,				
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,				
RA	Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,				
RA	Schrml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,				
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,				
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,				
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Cariboldi M.,				
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,				
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,				
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,				
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,				
RA	Suzuki H., Toyo-Oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,				
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,				
RA	Hayashizaki Y.;				
RL	*Functional annotation of a full-length mouse cDNA collection.*;				
RL	Nature 409:685-690(2001).				
DR	EMBL; AK018744; BAB31384.1; -.				

DR MGD; MGI:95832; Grn.  
DR InterPro; IPR000118; Granulin.  
DR InterPro; IPR001211; PhospholipaseA2.  
DR Pfam; PF00396; granulin; 7.  
DR SMART; SM00277; GRAN; 7.  
DR PROSITE; PS00799; GRANULINS; 7.  
DR PROSITE; PS00118; PA2\_HIS; UNKNOWN\_1.  
SQ SEQUENCE 589 AA; 63405 MW; 1DB8229CA13CA292 CRC64;

Query Match 100.0%; Score 69; DB 11; Length 589;  
Best Local Similarity 100.0%; Pred. No. 0.0003;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PDAKTQCPDDST 12  
|||||  
DB 208 PDAKTQCPDDST 219  
|||||

## RESULT 2

ID Q9QWB4 PRELIMINARY; PRT; 51 AA.  
AC Q9QWB4;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE PC cell-derived growth factor, PCDGF-EPITHELIN precursor homolog (Fragments).  
DE (Fragments).  
OS Mus sp.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10095;  
RN [1]  
RP SEQUENCE  
RX MEDLINE=93266526; PubMed=8496151;  
RA Zhou J., Gao G., Crabb J.W., Serrero G.;  
RT "Purification of an autocrine growth factor homologous with mouse epithelin precursor from a highly tumorigenic cell line.";  
RL J. Biol. Chem. 268:10863-10869(1993).  
FT NON\_TER 1  
FT NON\_TER 29  
FT NON\_TER 51  
FT NON\_TER 51  
SQ SEQUENCE 51 AA; 5473 MW; C349143BE4E942AF CRC64;

Query Match 94.2%; Score 65; DB 11; Length 51;  
Best Local Similarity 91.7%; Pred. No. 0.00012;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 PDAKTQCPDDST 12  
|||||  
DB 11 PDAKTQCPDDST 22  
|||||

## RESULT 3

ID Q8T9J3 PRELIMINARY; PRT; 384 AA.  
AC Q8T9J3;  
DT 01-JUN-2002 (TREMBLrel. 21, Created)  
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE GM08106p.  
GN CG13252.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephyroidae; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,  
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,  
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;

RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY069266; AAL39411.1;  
SQ SEQUENCE 384 AA; 42646 MW; 20AF3F1ACA95954D CRC64;

Query Match 68.1%; Score 47; DB 5; Length 384;  
Best Local Similarity 66.7%; Pred. No. 1.8;  
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 PDAKTQCPDDST 12  
|||||  
DB 45 PDAKTQCPDDST 56  
|||||

## RESULT 4

ID Q89001 PRELIMINARY; PRT; 511 AA.  
AC Q89001;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE (strain 83-HD-B2) glycoprotein.  
OS Vesicular stomatitis virus.  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Rhabdoviridae; Vesiculovirus.  
OX NCBI\_TaxID=11276;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90376442; PubMed=2168974;  
RA Bilseil P.A., Nichol S.T.;  
RT "Polymerase errors accumulating during natural evolution of the glycoprotein gene of vesicular stomatitis virus Indiana serotype RT isolates.";  
RL J. Virol. 64:4873-4883(1990).  
DR EMBL; M35218; AAA48398.1;  
DR InterPro; IPR001903; Rhabd\_glycop.  
DR Pfam; PF00974; Rhabd\_glycop; 1.  
SQ SEQUENCE 511 AA; 57501 MW; 921EF8029D610957 CRC64;

Query Match 65.2%; Score 45; DB 12; Length 511;  
Best Local Similarity 66.7%; Pred. No. 5.6;  
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 PDAKTQCPDDST 12  
|||||  
DB 427 PDAKTQCPDDST 438  
|||||

## RESULT 5

ID Q89002 PRELIMINARY; PRT; 511 AA.  
AC Q89002;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE (strain 85-ES-B1) glycoprotein.  
OS Vesicular stomatitis virus.  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Rhabdoviridae; Vesiculovirus.  
OX NCBI\_TaxID=11276;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90376442; PubMed=2168974;  
RA Bilseil P.A., Nichol S.T.;  
RT "Polymerase errors accumulating during natural evolution of the glycoprotein gene of vesicular stomatitis virus Indiana serotype RT isolates.";  
RL J. Virol. 64:4873-4883(1990).  
DR EMBL; M35220; AAA48390.1;  
DR InterPro; IPR001903; Rhabd\_glycop.  
DR Pfam; PF00974; Rhabd\_glycop; 1.  
SQ SEQUENCE 511 AA; 57519 MW; 93B338029EB50957 CRC64;

Query Match 65.2%; Score 45; DB 12; Length 511;

Best Local Similarity 66.7%; Pred. No. 5.6; Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 PDAKTCPPDDST 12

Db 427 PDATSQLPDDST 438

RESULT 6

ID Q89005 PRELIMINARY; PRT; 511 AA.

AC Q89005;

DT 01-NOV-1996 (TReMBLrel. 01, Created)

DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)

DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)

DE (strain 86-GM-B) glycoprotein.

OS Vesicular stomatitis virus.

OC Viruses; ssRNA negative-strand viruses; Mononegavirales;

OC Rhabdoviridae; Vesiculovirus.

OX NCBI\_TaxID=11276;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-90376442; PubMed-2168974;

RA Bilseil P.A., Nichol S.T.;

RT "Polymerase errors accumulating during natural evolution of the

RT glycoprotein gene of vesicular stomatitis virus Indiana serotype

RT isolates."

RL J. Virol. 64:4873-4883(1990).

DR EMBL; M35228; AAA48398.1; -

DR InterPro; IPR001903; Rhabd\_glycop.

DR Pfam; PF00974; Rhabd\_glycop; 1.

SQ SEQUENCE 511 AA; 57484 MW; 687FE74FD58861F1 CRC64;

Query Match 65.2%; Score 45; DB 12; Length 511;

Best Local Similarity 66.7%; Pred. No. 5.6;

Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 PDAKTCPPDDST 12

Db 427 PDATSQLPDDST 438

RESULT 7

ID Q89008 PRELIMINARY; PRT; 511 AA.

AC Q89008;

DT 01-NOV-1996 (TReMBLrel. 01, Created)

DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)

DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)

DE (strain 85-ES-B2) glycoprotein.

OS Vesicular stomatitis virus.

OC Viruses; ssRNA negative-strand viruses; Mononegavirales;

OC Rhabdoviridae; Vesiculovirus.

OX NCBI\_TaxID=11276;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-90376442; PubMed-2168974;

RA Bilseil P.A., Nichol S.T.;

RT "Polymerase errors accumulating during natural evolution of the

RT glycoprotein gene of vesicular stomatitis virus Indiana serotype

RT isolates."

RL J. Virol. 64:4873-4883(1990).

DR EMBL; M35232; AAA48402.1; -

DR InterPro; IPR001903; Rhabd\_glycop.

DR Pfam; PF00974; Rhabd\_glycop; 1.

SQ SEQUENCE 511 AA; 57518 MW; 93B338029EB50177 CRC64;

Query Match

Best Local Similarity

Matches 8; Conservative

1; Mismatches 3; Indels

0; Gaps 0;

QY 1 PDAKTCPPDDST 12

Db 427 PDATSQLPDDST 438

Db 427 PDATSQLPDDST 438

RESULT 8

ID Q89625 PRELIMINARY; PRT; 511 AA.

AC Q89625;

DT 01-NOV-1996 (TReMBLrel. 01, Created)

DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)

DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)

DE (strain 83-HD-B1)

OS Vesicular stomatitis virus.

OC Viruses; ssRNA negative-strand viruses; Mononegavirales;

OC Rhabdoviridae; Vesiculovirus.

OX NCBI\_TaxID=11276;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-90376442; PubMed-2168974;

RA Bilseil P.A., Nichol S.T.;

RT "Polymerase errors accumulating during natural evolution of the

RT glycoprotein gene of vesicular stomatitis virus Indiana serotype

RT isolates."

RL J. Virol. 64:4873-4883(1990).

DR EMBL; M35217; AAA48387.1; -

DR EMBL; M35216; AAA48386.1; -

DR InterPro; IPR001903; Rhabd\_glycop.

DR Pfam; PF00974; Rhabd\_glycop; 1.

SQ SEQUENCE 511 AA; 57491 MW; FFB3A673DBB50956 CRC64;

Query Match

Best Local Similarity

Matches 8; Conservative

1; Mismatches 3; Indels

0; Gaps 0;

QY 1 PDAKTCPPDDST 12

Db 427 PDATSQLPDDST 438

RESULT 9

ID Q89767 PRELIMINARY; PRT; 511 AA.

AC Q89767;

DT 01-NOV-1996 (TReMBLrel. 01, Created)

DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)

DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)

DE (strain 86-ES-B1)

OS Vesicular stomatitis virus.

OC Viruses; ssRNA negative-strand viruses; Mononegavirales;

OC Rhabdoviridae; Vesiculovirus.

OX NCBI\_TaxID=11276;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-90376442; PubMed-2168974;

RA Bilseil P.A., Nichol S.T.;

RT "Polymerase errors accumulating during natural evolution of the

RT glycoprotein gene of vesicular stomatitis virus Indiana serotype

RT isolates."

RL J. Virol. 64:4873-4883(1990).

DR EMBL; M35224; AAA48394.1; -

DR EMBL; M35221; AAA48391.1; -

DR EMBL; M35223; AAA48393.1; -

DR InterPro; IPR001903; Rhabd\_glycop.

DR Pfam; PF00974; Rhabd\_glycop; 1.

SQ SEQUENCE 511 AA; 57493 MW; 367C936F286E64E1 CRC64;

Query Match

Best Local Similarity

Matches 8; Conservative

1; Mismatches 3; Indels

0; Gaps 0;

QY 1 PDAKTCPPDDST 12

Db 427 PDATSQLPDDST 438

```

RESULT 10
Q89878 ID Q89878 PRELIMINARY; PRT; 511 AA.
AC Q89878;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE (strain 85-GM-B).
OS Vesicular stomatitis virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Vesiculovirus.
OX NCBI_TaxID=11276;
RN [1]
RP MEDLINE=90376442; PubMed=2168974;
RA Bilseil P.A., Nichol S.T.;
RT "Polymerase errors accumulating during natural evolution of the
RT glycoprotein gene of vesicular stomatitis virus Indiana serotype
RT isolates."
RL J. Virol. 64:4873-4883(1990).
DR EMBL; M35229; AAA48399.1; -
DR EMBL; M35222; AAA48392.1; -
DR InterPro; IPR001903; Rhabd_glycop.
DR Pfam; PF00974; Rhabd_glycop; 1.
SQ SEQUENCE 511 AA; 57475 MW; F8D3A773DBB50956 CRC64;

Query Match 55.2%; Score 45; DB 12; Length 511;
Best Local Similarity 66.7%; Pred. No. 5.6;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 PDAKTCPPDDST 12
   ||| : ||| |
Db 427 PDATSQLPDDST 438

RESULT 11
Q8WPK9 ID Q8WPK9 PRELIMINARY; PRT; 1557 AA.
AC Q8WPK9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Similar to silk protein (Fragment).
GN BAC001.24.
OS Oikopleura dioica.
OC Eukaryota; Metazoa; Chordata; Urochordata; Appendicularia;
OC Oikopleuridae; Oikopleura.
OX NCBI_TaxID=34765;
RN [1]
RP SEQUENCE FROM N.A.
RA Seo H.C., Kube M., Edvardsen R.B., Jensen M.F., Beck A., Spriet E.,
RA Gorsky G., Thompson E.W., Lehrach H., Reinhardt R., Chourrout D.,
RA "The marine chordate Oikopleura dioica has a miniature genome."
RL Science 0:0-0(2001).
DR EMBL; AF374376; AAL56451.1; -
DR NON_TER 1557 1557
FT SEQUENCE 1557 AA; 170296 MW; 57CCCC0FDF98DDD CRC64;

Query Match 65.2%; Score 45; DB 5; Length 1557;
Best Local Similarity 70.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 PDAKTCPPDD 10
   || |||||
Db 377 PDLNNQCPDD 386

RESULT 12
Q90362 ID Q90362 PRELIMINARY; PRT; 358 AA.
AC Q90362; Q90361;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)

```

```

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical 39.8 kDa protein T22H2.6.
GN T22H2.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Lennard N.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; T22H2.6A (SHOWN HERE) AND
CC T22H2.6B; ARE PRODUCED BY ALTERNATIVE SPLICING.
DR EMBL; Z81595; CAB54304.1; -
DR EMBL; Z81595; CAB54305.1; -
DR InterPro; IPR000118; Granulin.
DR Pfam; PF00396; granulin; 3.
DR SMART; SM00277; GRAN; 3.
DR PROSITE; PS00799; GRANULINS; UNKNOWN_2.
KW Alternative splicing; Hypothetical protein.
FT VARSPLIC 315 358
FT ACCEFTCCPAGCHVCVGGKCEKHAKTMTFRFNVKDDDE
FT DQI -> FNFPRSAVRTLAQAQPAITVLMVESAKTLKL
FT (IN ISOFORM T22H2.6B).
SQ SEQUENCE 358 AA; 39754 MW; 2AD5B8F9B70D1595 CRC64;

Query Match 62.3%; Score 43; DB 5; Length 358;
Best Local Similarity 63.6%; Pred. No. 8.8;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 DAKTCPPDDST 12
   ||: ||| |
Db 99 DAETECSDDET 109

RESULT 13
Q8ZTL3 ID Q8ZTL3 PRELIMINARY; PRT; 463 AA.
AC Q8ZTL3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE C-5 cytosine-specific DNA methylase.
GN PAE3201.
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX PubMed=11792869;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stettler K.O., Simon M.I.,
RA Miller J.H.;
RA "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum."
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR EMBL; AE009917; AAL64748.1; -
DR InterPro; IPR001525; C5_DNA_meth.
DR Pfam; PF00145; DNA_methylase; 1.
DR PRINTS; PR00105; C5METTRFRASE.
DR TIGRFS; TIGR00675; dcm; 1.
DR PROSITE; PS00094; C5_MTASE_1; 1.
KW Methyltransferase; Complete proteome.
SQ SEQUENCE 463 AA; 52202 MW; E63FF2F59A21B18C CRC64;

Query Match 60.9%; Score 42; DB 17; Length 463;
Best Local Similarity 66.7%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 PDAKTCPPDDST 12
   || | |||||
Db 333 PDGTTDPDDST 344

```

Db 39 DAQTVCPDGT 49

Search completed: July 8, 2003, 16:27:58  
Job time : 27.8462 secs

```
RESULT 14
Q8YD93
ID Q8YD93 PRELIMINARY; PRT; 506 AA.
AC Q8YD93;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Periplasmic dipeptide transport protein precursor.
GN BMEI10284.
OS Brucella melitensis.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Bruciellaceae; Brucella.
OX NCBI_TaxID=29459;
RN [1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=16M / ATCC 23456 / BIOTYPE 1;
RX MEDLINE=20020109; PubMed=11756688;
RA Ivanova N., Anderson I., Redkar R.J., Patra G., Mujar C., Los T.,
RA Delvechio V.G., Kapatral V., Bhattacharya A., Lykidis A., Reznik G.,
RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltzman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Haselkorn R., Kypides N., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis."
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
DR EMBL; AE009666; AAL53526.1; -.
DR InterPro: IPR000914; SBP_bac_5.
DR Pfam: PF00496; SBP_bac_5; 1.
DR Complete proteome.
KW SEQUENCE 506 AA; 55159 MW; C2F386ADBE79A999 CRC64;

Query Match 60.9%; Score 42; DB 16; Length 506;
Best Local Similarity 58.3%; Pred. No. 19;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 PDAKTCPPDDST 12
Db 143 PIATCECPDPT 154

RESULT 15
Q90ZD0
ID Q90ZD0 PRELIMINARY; PRT; 147 AA.
AC Q90ZD0;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Granulin 1.
OS Brachydanio rerio (zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RN
RP SEQUENCE FROM N.A.
RC Cadieux B., Bennett H.P.J.;
RT "Zebrafish granulin 1, 2, and hybrid cDNA sequences."
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF273479; AAK56708.1; -.
DR InterPro: IPR000118; Granulin.
DR InterPro: IPR002221; WAP.
DR Pfam: PF00396; granulin; 1.
DR PROSITE: PS00317; 4_DISULFIDE_CORE; UNKNOWN_1.
SQ SEQUENCE 147 AA; 16157 MW; 7A5D6864073A6CCC CRC64;

Query Match 59.4%; Score 41; DB 13; Length 147;
Best Local Similarity 63.6%; Pred. No. 7.9;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 DAKTCPPDDST 12
Db 143 PIATCECPDPT 154
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2003, 16:08:55 ; Search time 35 Seconds  
(without alignments)  
53.300 Million cell updates/sec

Title: US-09-824-647-5

Perfect score: 73  
Sequence: 1 SARGTKCLRKIPR 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	73	100.0	14	AAW85479	Mouse GP88 autocri
2	73	100.0	589	AA14327	Mouse epithelin pr
3	73	100.0	589	AAW85474	Mouse GP88 autocri
4	73	100.0	589	AAE20520	Mouse granulin/epi
5	65	89.0	589	AA14325	Rat epithelin prec
6	58	79.5	413	AA14325	Human protein sequ
7	58	79.5	593	AA14326	Human epithelin pr
8	58	79.5	621	AA14326	Human epithelin pr
9	53	72.6	14	AAW85481	Human GP88 autocri
10	53	72.6	593	AA14327	Granulin sequence.

11	53	72.6	593	20	AAW85475	Human GP88 autocri
12	53	72.6	593	23	AAE20521	Human granulin/epi
13	43	58.9	51	23	ABP34269	Human ORF3242 prot
14	40	54.8	76	22	AAU47116	Propionibacterium
15	40	54.8	85	22	AAU41864	Propionibacterium
16	40	54.8	255	20	AAW88397	Mouse neuro-growth
17	40	54.8	255	23	ABW72368	Murine protein 180
18	40	54.8	275	21	AAI52141	Mouse TANGO 125 (T
19	40	54.8	275	23	AAE21081	Mouse T125 (TANGO-
20	40	54.8	278	20	AAW88392	Mouse neuro-growth
21	40	54.8	278	22	AAE05356	Mouse Notch4-like
22	39	53.4	24	23	AAO17774	Human beta-defensi
23	39	53.4	27	23	AAO17773	Human beta-defensi
24	39	53.4	27	23	AAO17781	Human beta-defensi
25	39	53.4	40	23	AAO17766	Human beta-defensi
26	39	53.4	41	23	AAU09708	Human beta-defensi
27	39	53.4	45	21	AA110600	Human SAP-3 mature
28	39	53.4	45	23	AAO17767	Human beta-defensi
29	39	53.4	45	23	AAU09709	Human beta-defensi
30	39	53.4	67	20	AAI07244	Beta-defensin faml
31	39	53.4	67	21	AA110602	Human SAP-3 pre-pr
32	39	53.4	67	23	AAO17768	Human beta-defensi
33	39	53.4	67	23	AAU91016	Transplant media a
34	39	53.4	67	23	AAU91036	Transplant media a
35	39	53.4	67	23	AAU09707	Human beta-defensi
36	39	53.4	76	22	AAU91916	C glutamicum prote
37	39	53.4	94	23	ABW09594	Alpha-catenin 10.3
38	39	53.4	452	23	AAU79412	Human transcriptio
39	39	53.4	1123	22	ABW60739	Drosophila melanog
40	38.5	52.7	120	22	AAO00795	Human polypeptide
41	38	52.1	77	23	ABP33768	Propionibacterium
42	38	52.1	79	22	AAU56641	Propionibacterium
43	38	52.1	105	23	ABP04877	Human ORFX protein
44	38	52.1	154	22	AAU58583	Propionibacterium
45	38	52.1	821	15	AA146723	Arabidopsis thalia

ALIGNMENTS

RESULT 1

AAW85479  
ID AAW85479 standard; Peptide; 14 AA.

XX AAW85479;  
XX AC

XX 15-MAR-1999 (first entry)  
XX DT

XX DE Mouse GP88 autocrine growth factor antigenic peptide S14R.

XX KW GP88; granulin; epithelin; mouse; growth factor; autocrine; tumour;  
KW cancer; viral infection; antagonist; therapy; diagnosis; antigen;  
KW antibody.

XX OS Mus sp.  
XX PN WO9852607-A1.

XX PD 26-NOV-1998.

XX PF 22-MAY-1998; 98WO-US10555.

XX PR 16-DEC-1997; 97US-0991862.

XX PR 23-MAY-1997; 97US-0863079.

XX (SERR/) SERRERO G.

XX Serrero G;

XX WPI; 1999-045276/04.

XX Composition containing antagonist of growth factor GP88 - useful for  
PT treating cancer and viral diseases and also for diagnosing disease

PT from altered GP88 expression

XX Example 8; Page 45; 86pp; English.

XX This is the amino acid sequence of peptide S14R, comprising amino  
CC acid residues S562-R575 of murine GP88 (see AAW85474). GP88 is an 88  
CC kDa glycoprotein autocrine growth factor that is expressed in a  
CC tightly regulated manner in normal cells, is overexpressed and  
CC unregulated in highly tumorigenic cells derived from normal cells,  
CC and which acts as a stringently required growth stimulator for the  
CC tumorigenic cells. S14R was used to raise neutralising antibodies  
CC to GP88. Antagonists to GP88, such as anti-GP88 antibodies, are  
CC used to treat diseases associated with increased expression of  
CC GP88, particularly cancer but also viral infections. Anti-GP88  
CC antibodies can also be used as diagnostic reagents and to deliver  
CC toxins or other compounds to GP88-expressing cells.

XX Sequence 14 AA;

Query Match 100.0%; Score 73; DB 20; Length 14;

Best Local Similarity 100.0%; Pred. No. 2.4e-06; Mismatches 0; Indels 0; Gaps 0;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SARGTKCLRKIKIPR 14

DB 1 SARGTKCLRKIKIPR 14

RESULT 2

ID AAR14327  
AAAR14327: standard; Protein; 589 AA.

XX AAR14327;

DT 17-JAN-1992 (first entry)

XX Mouse epithelin precursor.

DE ET; growth regulation; inhibition; stimulation.

KW Mus musculus.

XX Key

Location/Qualifiers

1..589

/label= precursor

/note= "claim 21, page 55"

280..335

/label= EP-1

/note= "claim 22, page 55"

205..261

/label= EP-2

/note= "claim 23, page 55"

59..114

/label= EP

/note= "claim 24, page 55"

123..179

/label= EP

/note= "claim 25, page 55"

362..416

/label= EP

/note= "claim 26, page 56"

440..495

/label= EP

/note= "claim 27, page 56"

515..570

/label= EP

/note= "claim 28, page 56"

XX WO9115510-A.

PN 17-OCT-1991.

XX 03-APR-1991; 91WO-US02321.

XX 13-MAR-1991; 91US-0083796.  
PR 03-APR-1990; 90US-0504508.

XX (BRIM ) BRISTOL-MYERS SQUIB.

XX Shoyab M., Plowman GD;

XX WPI; 1991-325168/44.

XX N-PSDB; AAO14340.

XX New cysteine-rich growth modulating proteins, epithelins - useful

PT as inhibitors of neoplastic cell growth and to promote wound

PT healing and treat psoriasis

XX Disclosure; Fig 23; 97pp; English.

XX ET-1 is a bifunctional growth regulator, capable of stimulating

CC the growth of some cell types while inhibiting the growth of others.

CC ET-2 is functionally similar to ET-1 w.r.t. growth inhibitory

CC bioactivity. In contrast, however, ET-2 is apparently not capable of

CC eliciting the growth stimulatory activity characteristic of ET-1 and,

CC in fact, antagonises this ET-1 activity.

CC See also AAO14338-40, AAO14952-53, AAR14328-9 and AAR15315-20.

XX Sequence 589 AA;

Query Match 100.0%; Score 73; DB 12; Length 589;

Best Local Similarity 100.0%; Pred. No. 9.8e-05;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SARGTKCLRKIKIPR 14

DB 562 SARGTKCLRKIKIPR 575

RESULT 3

AAW85474

ID AAW85474 standard; Protein; 589 AA.

XX AAW85474;

XX 15-MAR-1999 (first entry)

XX Mouse GP88 autocrine growth factor.

DE GP88; granulin; epithelin; mouse; growth factor; autocrine; tumour;

KW cancer; viral infection; antagonist; therapy; diagnosis.

XX Mus sp.

XX Key

Location/Qualifiers

8

/note= "encoded by CTG"

Misc-difference 54

/note= "encoded by AGC"

Misc-difference 377

/note= "encoded by TGA"

Peptide

208..219

/note= "P12T peptide used to raise antibody".

Peptide

344..362

/note= "K19T peptide, used to raise antibody"

Peptide

562..575

/note= "S14R peptide, used to raise antibody"

XX WO9852607-A1.

PN 26-NOV-1998.

XX 22-MAY-1998; 98WO-US10555.

XX 16-DEC-1997; 97US-0991862.

XX 23-MAY-1997; 97US-0863079.



XX PA (SERR/) SERRERO G.  
 XX PI Serrero G;  
 XX XX  
 XX XX  
 DR WPI; 1999-045276/04.  
 DR N-PSDB; AAV82824.  
 XX  
 XX Composition containing antagonist of growth factor GP88 - useful for  
 PT treating cancer and viral diseases and also for diagnosing disease  
 PT from altered GP88 expression  
 XX  
 XX Example 5; Fig 8A-D; 86pp; English.  
 PS  
 XX This is the amino acid sequence of murine GP88, an 88 kDa  
 CC glycoprotein autocrine growth factor and epithelin/granulin  
 CC precursor that is expressed in a tightly regulated manner in normal  
 CC cells, is overexpressed and unregulated in highly tumorigenic cells  
 CC derived from normal cells, and which acts as a stringently required  
 CC growth stimulator for the tumorigenic cells. Inhibition of GP88  
 CC expression or action in the tumorigenic cells results in an  
 CC inhibition of the tumorigenic properties of the overproducing  
 CC cells. Murine GP88 cDNA (see AAV82824) was obtained from the highly  
 CC tumorigenic PC cell line. Antagonists to GP88 are used to treat  
 CC diseases associated with increased expression of GP88, particularly  
 CC cancer but also viral infections. Fragments of GP88 are used to  
 CC raise specific antibodies (used as antagonists, as diagnostic  
 CC reagents and for delivering toxins or other compounds to GP88-  
 CC expressing cells) and to screen for antibodies. Methods are  
 CC provided for diagnosing disease, or determining susceptibility to  
 CC disease, resulting from altered GP88 activity.  
 XX  
 XX Sequence 589 AA;  
 SQ  
 Query Match 100.0%; Score 73; DB 20; Length 589;  
 Best Local Similarity 100.0%; Pred. No. 9.8e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SARGTKCLRRKKIPR 14  
 Db 562 SARGTKCLRRKKIPR 575  
 RESULT 4  
 AAE20520  
 ID AAE20520 standard; Protein; 589 AA.  
 XX AC AAE20520;  
 XX  
 XX 01-JUL-2002 (first entry)  
 DT  
 XX Mouse granulin/epithelin precursor (GP88) protein.  
 DE  
 XX Mouse; granulin precursor; GP88; cytostatic; tumorigenicity; tamoxifen;  
 KW antineoplastic; antioestrogen therapy; skin cancer.  
 XX  
 XX Mus sp.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FH Misc-difference 335  
 FT /label= Unknown  
 FT /note= "Encoded by ATG"  
 FT Region 344..362  
 FT /note= "Regions used as antigens to raise anti-GP88  
 FT antibodies"  
 FT Region 562..575  
 FT /note= "Regions used as antigens to raise anti-GP88  
 FT antibodies"  
 FT Misc-difference 586  
 FT /note= "Encoded by ACA"  
 XX  
 XX US2002025543-A1.  
 PN  
 XX

PD 28-FEB-2002.  
 XX  
 XX 15-JUN-2001; 2001US-0880842.  
 XX PF  
 XX 23-MAY-1997; 97US-0863079.  
 PR 08-DEC-1999; 99US-0456886.  
 XX  
 XX (SERR/) SERRERO G.  
 PA  
 XX Serrero G;  
 PI  
 XX WPI; 2002-267529/31.  
 DR N-PSDB; AAD32849.  
 DR  
 XX  
 XX Diagnosing tumorigenicity in a human, comprising obtaining a cell  
 PT sample, detecting GP88 in the cells, and determining the number of GP88  
 PT positive cells in the sample -  
 XX  
 XX Disclosure; Fig 8; 50pp; English.  
 CC  
 CC The invention relates to a method for diagnosing tumorigenicity in a  
 CC human. The method comprises obtaining a biological sample containing  
 CC cells from the patient, detecting GP88 in the cells of the sample,  
 CC and determining the number of GP88 positive cells in the sample, and  
 CC determining the ratio of GP88 positive cells to the total number of cells  
 CC in the sample. The invention also relates to a method for determining  
 CC if a human patient is resistant to the antineoplastic effects of  
 CC antioestrogen therapy. The method is useful for diagnosing  
 CC tumorigenicity in a sample, such as blood, serum, plasma, urine, nipple  
 CC aspirate, cerebrospinal fluid, liver, kidney, breast, bone, brain, lung,  
 CC colon, or skin cancer. The method can be used to treat or prevent  
 CC re-occurrence of cancer in a patient, by administering tamoxifen if the  
 CC sample contains less than 10 % GP88, or less than 5 % GP88 positive  
 CC cells. The present sequence is mouse granulin/epithelin precursor (GP88)  
 CC protein.  
 XX  
 XX Sequence 589 AA;  
 SQ  
 Query Match 100.0%; Score 73; DB 23; Length 589;  
 Best Local Similarity 100.0%; Pred. No. 9.8e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SARGTKCLRRKKIPR 14  
 Db 562 SARGTKCLRRKKIPR 575  
 RESULT 5  
 AARI4325  
 ID AARI4325 standard; Protein; 589 AA.  
 XX AC AARI4325;  
 XX  
 XX 17-JAN-1992 (first entry)  
 DT  
 XX Rat epithelin precursor.  
 DE  
 XX ET; growth regulation; inhibition; stimulation.  
 XX  
 XX Rattus rattus.  
 OS  
 XX Key Location/Qualifiers  
 FH Protein 1..589  
 FT /label= precursor  
 FT /note= "Claim 11, page 54"  
 FT 280..335  
 FT /label= EP-1  
 FT /note= "claim 12, page 54"  
 FT 205..261  
 FT /label= EP-2  
 FT /note= "claim 13, page 54"  
 FT 59..114  
 FT /label= EP  
 FT Peptide

FT /note= "claim 14, page 54"  
FT 123..179  
FT /label= EP  
FT /note= "claim 15, page 54"  
FT 362..416  
FT /label= EP  
FT /note= "claim 16, page 54"  
FT 440..495  
FT /label= EP  
FT /note= "claim 17, page 54"  
FT 515..570  
FT /label= EP  
FT /note= "claim 18, page 55"  
FT  
FT  
PN W09115510-A.  
XX  
XX  
PD 17-OCT-1991.  
XX  
XX 03-APR-1991; 91WO-US02321.  
XX  
PR 13-MAR-1991; 91US-0083796.  
PR 03-APR-1990; 90US-0504508.  
XX  
PA (BRIM ) BRISTOL-MYERS SQUIB.  
XX  
XX Shoyab M, Plowman GD;  
PI  
XX WPI; 1991-325168/44.  
DR N-PSDB; AAQ14338.  
XX  
XX New cysteine-rich growth modulating proteins, epithelins - useful  
PT as inhibitors of neoplastic cell growth and to promote wound  
PT healing and treat psoriasis  
XX  
PS Disclosure; Fig 18; 97pp; English.  
XX  
CC ET-1 is a bifunctional growth regulator, capable of stimulating  
CC the growth of some cell types while inhibiting the growth of others.  
CC ET-2 is functionally similar to ET-1 w.r.t. growth inhibitory  
CC bioactivity. In contrast, however, ET-2 is apparently not capable of  
CC eliciting the growth stimulatory activity characteristic of ET-1 and,  
CC in fact, antagonises this ET-1 activity.  
CC See also AAQ14338-40, AAQ14952-53, AAR14328-9 and AAR15315-20.  
XX  
SQ Sequence 589 AA;  
Query Match 89.0%; Score 65; DB 12; Length 589;  
Best Local Similarity 85.7%; Pred. No. 0.0027;  
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 SARGTKCLRRKKIPR 14  
DB 562 SARGTKCLRRKKTPR 575  
II:|||||IIIIII  
RESULT 6  
AAB94550  
ID AAB94550 standard; Protein; 413 AA.  
AC AAB94550;  
XX  
XX 26-JUN-2001 (first entry)  
XX Human protein sequence SEQ ID NO:15310.  
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.  
XX Homo sapiens.  
XX EP1074617-A2.  
PN  
XX  
XX 07-FEB-2001.  
PD  
XX

28-JUL-2000; 2000EP-0116126.  
XX  
XX 29-JUL-1999; 99JP-0248036.  
PR 27-AUG-1999; 99JP-0300253.  
PR 11-JAN-2000; 2000JP-0118776.  
PR 02-MAY-2000; 2000JP-0183767.  
PR 09-JUN-2000; 2000JP-0241899.  
XX  
XX (HELI-) HELIX RES INST.  
XX  
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
XX  
XX WPI; 2001-318749/34.  
XX  
PT Primer sets for synthesizing polynucleotides, particularly the 5602  
PT full-length cDNAs defined in the specification, and for the detection  
PT and/or diagnosis of the abnormality of the proteins encoded by the  
PT full-length cDNAs -  
XX  
XX Claim 8; SEQ ID 15310; 2537pp + CD ROM; English.  
XX  
CC The present invention describes primer sets for synthesising 5602  
CC full-length cDNAs defined in the specification. Where a primer set  
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
CC to the complementary strand of a polynucleotide which comprises one of  
CC the 5602 nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in  
CC the specification. The primer sets can be used in antisense therapy and  
CC in gene therapy. The primers are useful for synthesising polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAR13628 and  
CC AAR13633 to AAR18742 represent human cDNA sequences; AAB92446 to  
CC AAB93893 represent human amino acid sequences; and AAR13629 to AAR13632  
CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.  
XX  
SQ Sequence 413 AA;  
Query Match 79.5%; Score 58; DB 22; Length 413;  
Best Local Similarity 71.4%; Pred. No. 0.035;  
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 1 SARGTKCLRRKKIPR 14  
DB 386 AARGTKCLRRRAPR 399  
II:|||||IIIIIIIIII  
RESULT 7  
AAR14326  
ID AAR14326 standard; Protein; 593 AA.  
XX  
XX AAR14326;  
XX  
XX 17-JAN-1992 (first entry)  
XX Human epithelin precursor.  
XX  
XX ET; growth regulation; inhibition; stimulation.  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
XX Protein 1..593  
XX /label= precursor  
FT

FT Protein /note- "claim 1, page 53"  
 FT 282..337  
 FT /label- EP-1  
 FT /note- "claim 2, page 53"  
 FT 206..262  
 FT Protein /label- EP-2  
 FT /note- "claim 3, page 53"  
 FT 59..114  
 FT Peptide /label- EP  
 FT /note- "claim 4, page 53"  
 FT 124..180  
 FT Peptide /label- EP  
 FT /note- "claim 5, page 53"  
 FT 364..418  
 FT Peptide /label- EP  
 FT /note- "claim 6, page 53"  
 FT 442..497  
 FT Peptide /label- EP  
 FT /note- "claim 7, page 53"  
 FT 519..574  
 FT Peptide /label- EP  
 FT /note- "claim 8, page 55"  
 FT  
 XX WO9115510-A.  
 PN  
 XX 17-OCT-1991.  
 PD  
 XX  
 XX 03-APR-1991; 91WO-US02321.  
 PF  
 XX  
 XX 13-MAR-1991; 91US-0083796.  
 PR  
 XX 03-APR-1990; 90US-0504508.  
 XX  
 XX (BRIM ) BRISTOL-MYERS SQUIB.  
 PA  
 XX  
 XX Shoyab M, Plowman GD;  
 PI  
 XX WPI; 1991-325168/44.  
 DR  
 XX N-PSDB; AAQ14339.  
 XX  
 XX New cysteine-rich growth modulating proteins, epithelins - useful  
 PT as inhibitors of neoplastic cell growth and to promote wound  
 PT healing and treat psoriasis  
 PT  
 XX Disclosure; Fig 22; 97pp; English.  
 PS  
 XX ET-1 is a bifunctional growth regulator, capable of stimulating  
 CC the growth of some cell types while inhibiting the growth of others.  
 CC ET-2 is functionally similar to ET-1 w.r.t. growth inhibitory  
 CC bioactivity. In contrast, however, ET-2 is apparently not capable of  
 CC eliciting the growth stimulatory activity characteristic of ET-1 and,  
 CC in fact, antagonises this ET-1 activity.  
 CC See also AAQ14338-40, AAQ14952-53, AAR14328-9 and AAR15315-20.  
 XX  
 SQ Sequence 593 AA;

Query Match 79.5%; Score 58; DB 12; Length 593;  
 Best Local Similarity 71.4%; Pred. No. 0.05;  
 Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 SARGTKCLRKIKIPR 14  
 DB 566 AARGTKCLRREAPR 579

RESULT 8  
 AAB43971  
 ID AAB43971 standard; Protein; 621 AA.  
 XX  
 XX AAB43971;  
 XX  
 XX 08-FEB-2001 (first entry)  
 XX  
 XX Human cancer associated protein sequence SEQ ID NO:1416.

XX Human; cancer associated gene; cancer antigen; detection; cancer;  
 KW diagnosis; cytostatic; proliferative; vulnery; immunomodulator;  
 KW antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;  
 KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;  
 KW dermatological; neuroprotective; thrombolytic; coagulant; nootropic;  
 KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;  
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;  
 KW allergic reaction; graft versus host disease; organ rejection;  
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;  
 KW neurological disease; drug screening.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200055350-A1.  
 PN  
 XX 21-SEP-2000.  
 PD  
 XX  
 XX 08-MAR-2000; 2000WO-US05882.  
 PF  
 XX  
 XX 12-MAR-1999; 99US-0124270.  
 PR  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 XX  
 XX Rosen CA, Ruben SM;  
 PI  
 XX  
 XX WPI; 2000-587533/55.  
 DR  
 XX N-PSDB; AAC78180.  
 DR  
 XX  
 XX Novel isolated nucleic acids comprising sequences encoding peptides  
 PT useful for treating or diagnosing e.g. cancer -  
 XX  
 XX Claim 11; Page 2094-2096; 2352pp; English.  
 PS  
 XX  
 XX AAC77607 to AAC78448 encode the human cancer associated proteins given  
 CC in AAB43398 to AAB44239. The proteins can have activities based on the  
 CC tissues and cells the genes are expressed in. Example of activities  
 CC include: cytostatic; proliferative; vulnery; immunomodulator;  
 CC antidiabetic; antiasthmatic; antirheumatic; antiarthritic;  
 CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;  
 CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;  
 CC nootropic; vasotropic; antipsoriatic and antiangiogenic. The  
 CC polynucleotides and polypeptides can be used for preventing, treating or  
 CC ameliorating medical conditions and diagnosing pathological conditions.  
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from  
 CC the present invention may be used to treat immune disorders by activating  
 CC or inhibiting the proliferation, differentiation or mobilisation of  
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune  
 CC disorders, allergic reactions, graft versus host disease and organ  
 CC rejection, modulate haemostatic or thrombolytic activity, modulate  
 CC inflammation, cancers, cardiovascular disorders, neurological disease and  
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,  
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to  
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of  
 CC the present invention.

Sequence 621 AA;  
 Query Match 79.5%; Score 58; DB 21; Length 621;  
 Best Local Similarity 71.4%; Pred. No. 0.052;  
 Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 SARGTKCLRKIKIPR 14  
 DB 594 AARGTKCLRREAPR 607

RESULT 9  
 AAW85481  
 ID AAW85481 standard; Peptide; 14 AA.  
 XX  
 XX AAW85481;  
 AC  
 XX

DT 15-MAR-1999 (first entry)  
 XX Human GP88 autocrine growth factor antigenic peptide A14R.  
 DE GP88; granulatin; epithillin; human; growth factor; autocrine; tumour;  
 KW cancer; viral infection; antagonist; therapy; diagnosis; antigen;  
 KW antibody.  
 XX Homo sapiens.  
 OS WO9852607-A1.  
 PN 26-NOV-1998.  
 XX 22-MAY-1998; 98WO-US10555.  
 XX 16-DEC-1997; 97US-0991862.  
 PR 23-MAY-1997; 97US-0863079.  
 XX (SERR/) SERRERO G.  
 PA Serrero G;  
 PI WPI; 1999-045276/04.  
 DR Composition containing antagonist of growth factor GP88 - useful for  
 PT treating cancer and viral diseases and also for diagnosing disease  
 PT from altered GP88 expression  
 XX Example 8; Page 45; 86pp; English.  
 PS This is the amino acid sequence of peptide A14R, comprising amino  
 CC acid residues A566-R579 of human GP88 (see AAW85475). GP88 is an 88  
 CC kDa glycoprotein autocrine growth factor that is expressed in a  
 CC tightly regulated manner in normal cells, is overexpressed and  
 CC unregulated in highly tumorigenic cells derived from normal cells,  
 CC and which acts as a stringently required growth stimulator for the  
 CC tumorigenic cells. A14R was used in an attempt to raise  
 CC neutralising antibodies to GP88 (see also AAW85480). Antagonists to  
 CC GP88, such as anti-GP88 antibodies, are used to treat diseases  
 CC associated with increased expression of GP88, particularly cancer  
 CC but also viral infections. Anti-GP88 antibodies can also be used  
 CC as diagnostic reagents and to deliver toxins or other compounds to  
 CC GP88-expressing cells.  
 XX Sequence 14 AA;  
 SQ Query Match 72.6%; Score 53; DB 20; Length 14;  
 Best Local Similarity 75.0%; Pred. No. 0.0098;  
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 RGTCLRKIKIPR 14  
 DB |||||::: ||  
 3 RGTCLRREAPR 14  
 RESULT 10  
 AAR48673  
 ID AAR48673 standard; Protein; 593 AA.  
 XX AAR48673;  
 AC 22-APR-1994 (first entry)  
 DT Granulin sequence.  
 DE Granulin; keratinocytes; wound healing; inhibition; peptide;  
 KW granulocytes; leucocytes.  
 KW Homo sapiens.  
 OS Key Location/Qualifiers  
 XX Serrero G;  
 FH WPI; 1999-045276/04.  
 FT Misc-difference 452

FT Misc-difference 539 /note- "Valine encoded by ATG."  
 FT /note- "Glycine encoded by CAG."  
 XX WO9315195-A.  
 XX 05-AUG-1993.  
 PD 28-FEB-1992; 92WO-CA00089.  
 XX 03-FEB-1992; 92US-0829233.  
 PR (SOLO/) SOLOMON S.  
 XX Solomon S;  
 PI WPI; 1993-320328/40.  
 DR N-PSDB; AAO49052.  
 XX New cysteine rich granulin peptide(s) from leucocyte(s) - are  
 PT keratinocyte inhibitors useful topically for wound healing  
 XX Disclosure; Figure 4c; 53pp; English.  
 PS The granulin inhibits keratinocytes and is useful in formulations  
 CC for promoting the healing of wounds.  
 CC Sequence 593 AA;  
 SQ Query Match 72.6%; Score 53; DB 14; Length 593;  
 Best Local Similarity 75.0%; Pred. No. 0.4;  
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 RGTCLRKIKIPR 14  
 DB |||||::: ||  
 568 RGTCLRREAPR 579  
 RESULT 11  
 AAW85475  
 ID AAW85475 standard; Protein; 593 AA.  
 XX AAW85475;  
 AC 15-MAR-1999 (first entry)  
 DT Human GP88 autocrine growth factor.  
 XX GP88; granulatin; epithillin; human; growth factor; autocrine; tumour;  
 KW cancer; viral infection; antagonist; therapy; diagnosis.  
 XX Homo sapiens.  
 OS Key Location/Qualifiers  
 FH Peptide 340..364  
 FT /note- "E19V peptide used to raise antibody"  
 FT Peptide 566..579  
 FT /note- "A14R peptide used to raise antibody"  
 XX WO9852607-A1.  
 PN 26-NOV-1998.  
 XX 22-MAY-1998; 98WO-US10555.  
 PR 16-DEC-1997; 97US-0991862.  
 PR 23-MAY-1997; 97US-0863079.  
 XX (SERR/) SERRERO G.  
 PA Serrero G;  
 XX WPI; 1999-045276/04.  
 DR Misc-difference 452

DR N-PSDB; AAV82825.

XX Composition containing antagonist of growth factor GP88 - useful for

PT treating cancer and viral diseases and also for diagnosing disease

PT from altered GP88 expression

XX

PS Example 5; Fig 9A; 86pp; English.

PS

XX This is the amino acid sequence of human GP88, an 88 kDa

CC glycoprotein autocrine growth factor and epithelin/granulin

CC precursor that is expressed in a tightly regulated manner in normal

CC cells, is overexpressed and unregulated in highly tumorigenic cells

CC derived from normal cells, and which acts as a stringently required

CC growth stimulator for the tumorigenic cells. Inhibition of GP88

CC expression or action in the tumorigenic cells results in an

CC inhibition of the tumorigenic properties of the overproducing

CC cells. Antagonists to GP88 are used to treat diseases associated

CC with increased expression of GP88, particularly cancer but also

CC viral infections. Fragments of GP88 are used to raise specific

CC antibodies (used as antagonists, as diagnostic reagents and for

CC delivering toxins or other compounds to GP88-expressing cells) and

CC to screen for antibodies. Methods are provided for diagnosing

CC disease, or determining susceptibility to disease, resulting from

CC altered GP88 activity.

XX

SQ Sequence 593 AA;

Query Match 72.6%; Score 53; DB 20; Length 593;

Best Local Similarity 75.0%; Pred. No. 0.4;

Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RGTCLRLKKIPR 14

DB 568 RGTCLRLREAPR 579

|||||::: ||

RESULT 12

AAE20521

ID AAE20521 standard; Protein; 593 AA.

XX

AC AAE20521;

XX

DT 01-JUL-2002 (first entry)

XX

DE Human granulin/epithelin precursor (GP88) protein.

XX

KW Human; granulin precursor; GP88; cytostatic; tumorigenicity; tamoxifen;

KW antineoplastic; antioestrogen therapy; skin cancer.

OS

OS Homo sapiens.

XX

Key Location/Qualifiers

FT Misc-difference 273

FT /note= "Encoded by AAG"

FT 346..364

FT /note= "Region used as antigen to develop anti-human

FT GP88 neutralising antibody"

XX

US2002025543-A1.

XX

PD 28-FEB-2002.

XX

PD 15-JUN-2001; 2001US-0880842.

XX

PR 23-MAY-1997; 97US-0863079.

PR 08-DEC-1999; 99US-0456886.

XX

PA (SERR/) SERRERO G.

XX

PI Serrero G;

XX

DR WPI; 2002-267529/31.

DR N-PSDB; AAD32850.

XX Diagnosing tumorigenicity in a human, comprising obtaining a cell

PT sample, detecting GP88 in the cells, and determining the number of GP88

PT positive cells in the sample -

XX

PS Disclosure; Fig 9B; 50pp; English.

XX

CC The invention relates to a method for diagnosing tumorigenicity in a

CC human. The method comprises obtaining a biological sample containing

CC cells from the patient, detecting GP88 in the cells of the sample, and

CC determining the number of GP88 positive cells in the sample, and

CC determining the ratio of GP88 positive cells to the total number of cells

CC in the sample. The invention also relates to a method for determining

CC if a human patient is resistant to the antineoplastic effects of

CC antioestrogen therapy. The method is useful for diagnosing

CC tumorigenicity in a sample, such as blood, serum, plasma, urine, nipple

CC aspirate, cerebrospinal fluid, liver, kidney, breast, bone, brain, lung,

CC colon, or skin cancer. The method can be used to treat or prevent

CC re-occurrence of cancer in a patient, by administering tamoxifen if the

CC sample contains less than 10 % GP88, or less than 5 % GP88 positive

CC cells. The present sequence is human granulin/epithelin precursor (GP88)

CC protein.

XX

SQ Sequence 593 AA;

Query Match 72.6%; Score 53; DB 23; Length 593;

Best Local Similarity 75.0%; Pred. No. 0.4;

Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RGTCLRLKKIPR 14

DB 568 RGTCLRLREAPR 579

|||||::: ||

RESULT 13

ABP34269

ID ABP34269 standard; Protein; 51 AA.

XX

AC ABP34269;

XX

DT 08-JUL-2002 (first entry)

XX

DE Human ORF3242 protein, SEQ ID NO:6484.

XX

KW Human; ORF; open reading frame; ORFX; drug screening; diagnosis;

KW disease monitoring; cytokine; cell proliferation; cell differentiation;

KW immune modulation; haematopoiesis regulation; tissue growth;

KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;

KW thrombolytic; tumour inhibition; bodily characteristic; fertility;

KW behaviour; cancer; proliferative disorder; neurological disorder;

KW cardiovascular disease; immune system disorder; organ transplantation;

KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;

KW hypothyroidism; cholesterol ester storage disease; infection; vulnery;

KW vasotropic; antipsoriatic; antidiabetic; cytostatic; neurotropic;

KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;

KW cardiatic; hypotensive; antithyroid; antiinflammatory; immunomodulator;

KW dermatological; analgesic; virucide; antibacterial; fungicide.

XX

OS Homo sapiens.

XX

PN WO200190366-A2.

XX

PD 29-NOV-2001.

XX

PF 24-MAY-2001; 2001WO-US17076.

XX

PR 24-MAY-2000; 2000US-206690P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Leach MD, Shimkets RA;

XX

DR WPI; 2002-106200/14.



XX 01-NOV-2001.  
XX  
XX 20-APR-2001; 2001WO-US12865.  
XX  
XX 21-APR-2000; 2000US-199047P.  
XX  
XX 02-JUN-2000; 2000US-208841P.  
XX  
XX 07-JUL-2000; 2000US-216747P.  
XX  
XX (CORI-) CORTXA CORP.  
XX  
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
XX L'maisonneuve J, Zhang Y, Jen S, Carter D;  
XX  
XX WPI; 2001-616774/71.  
XX  
XX N-PSDB; AAS59515.  
XX  
XX Propionibacterium acnes polypeptides and nucleic acids useful for  
XX vaccinating against and diagnosing infections, especially useful for  
XX treating acne vulgaris  
XX  
XX Example 1; SEQ ID No 3059; 1069pp; English.  
XX  
XX Sequences AAU39105-AAUG68017 represent Propionibacterium acnes immunogenic  
XX polypeptides. The proteins and their associated DNA sequences are used in  
XX the treatment, prevention and diagnosis of medical conditions caused by  
XX P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
XX pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.  
XX P. acnes is also involved in infections of bone, joints and the central  
XX nervous system, however it is particularly involved in the inflammatory  
XX lesions associated with acne vulgaris. A method for detecting the  
XX presence or absence of P. acnes in a patient comprises contacting a  
XX sample with a binding agent that binds to the proteins of the invention  
XX and determining the amount of bound protein in the sample. The  
XX polypeptides may be used as antigens in the production of antibodies  
XX specific for P. acnes proteins. These antibodies can be used to  
XX downregulate expression and activity of P. acnes polypeptides and  
XX therefore treat P. acnes infections. The antibodies may also be used as  
XX diagnostic agents for determining P. acnes presence, for example, by  
XX enzyme linked immunosorbent assay (ELISA).  
XX Note: The sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 85 AA;

Query Match 54.8%; Score 40; DB 22; Length 85;  
Best Local Similarity 53.8%; Pred. No. 13;  
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 ARGTKCLRKKIPR 14  
||| ||| :: |  
Db 57 ARGVCLRLSRVAR 69

Search completed: July 8, 2003, 16:24:15  
Job time : 36 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2003, 16:21:00 ; Search time 8.61539 Seconds  
(without alignments)  
47.812 Million cell updates/sec

Title: US-09-824-647-5  
Perfect score: 73  
Sequence: 1 SARGTKLRKKIPR 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2.6/pdata/1/1aa/5A-COMB.pep:\*  
2: /cgn2.6/pdata/1/1aa/5B-COMB.pep:\*  
3: /cgn2.6/pdata/1/1aa/6A-COMB.pep:\*  
4: /cgn2.6/pdata/1/1aa/6B-COMB.pep:\*  
5: /cgn2.6/pdata/1/1aa/PCTUS-COMB.pep:\*  
6: /cgn2.6/pdata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	73	100.0	14	4 US-08-991-862-5	Sequence 5, Appl
2	73	100.0	589	1 US-07-668-648-6	Sequence 6, Appl
3	73	100.0	589	2 US-08-429-998-6	Sequence 6, Appl
4	73	100.0	589	2 US-08-431-333-6	Sequence 6, Appl
5	73	100.0	589	4 US-08-991-862-2	Sequence 2, Appl
6	73	100.0	589	5 PCT-US91-02321-6	Sequence 6, Appl
7	65	89.0	589	1 US-07-668-648-2	Sequence 2, Appl
8	65	89.0	589	2 US-08-429-998-2	Sequence 2, Appl
9	65	89.0	589	2 US-08-431-333-2	Sequence 2, Appl
10	65	89.0	589	5 PCT-US91-02321-2	Sequence 2, Appl
11	58	79.5	593	1 US-07-668-648-4	Sequence 4, Appl
12	58	79.5	593	2 US-08-429-998-4	Sequence 4, Appl
13	58	79.5	593	2 US-08-431-333-4	Sequence 4, Appl
14	58	79.5	593	5 PCT-US91-02321-4	Sequence 4, Appl
15	53	72.6	14	4 US-08-991-862-7	Sequence 7, Appl
16	53	72.6	593	4 US-08-991-862-17	Sequence 17, Appl
17	40	54.8	278	4 US-09-724-864-52	Sequence 52, Appl
18	38	52.1	263	3 US-09-035-706-5	Sequence 5, Appl
19	38	52.1	263	3 US-08-955-841-5	Sequence 5, Appl
20	38	52.1	263	4 US-09-390-425-5	Sequence 5, Appl
21	38	52.1	263	4 US-09-566-906-5	Sequence 5, Appl
22	38	52.1	821	1 US-07-928-464-2	Sequence 2, Appl
23	38	52.1	821	1 US-08-003-311B-2	Sequence 2, Appl
24	38	52.1	821	1 US-08-261-432-2	Sequence 2, Appl
25	38	52.1	821	5 PCT-US93-07347-2	Sequence 2, Appl
26	36	49.3	520	4 US-09-326-203A-2	Sequence 2, Appl
27	36	49.3	598	2 US-08-937-540-2	Sequence 2, Appl

28	36	49.3	598	4	US-09-398-395A-26	Sequence 26, Appl
29	35	47.9	41	1	US-08-112-208C-7	Sequence 7, Appl
30	35	47.9	41	1	US-08-248-819A-7	Sequence 7, Appl
31	35	47.9	41	2	US-08-337-646A-7	Sequence 7, Appl
32	35	47.9	41	2	US-08-856-531-7	Sequence 7, Appl
33	35	47.9	41	2	US-08-856-034-7	Sequence 7, Appl
34	35	47.9	41	4	US-08-927-326-7	Sequence 7, Appl
35	34	46.6	76	4	US-09-309-487-16	Sequence 16, Appl
36	34	46.6	92	4	US-09-309-487-22	Sequence 22, Appl
37	34	46.6	301	1	US-09-230-421-2	Sequence 2, Appl
38	34	46.6	314	1	US-08-589-446-6	Sequence 6, Appl
39	34	46.6	314	1	US-08-444-882-6	Sequence 6, Appl
40	34	46.6	314	2	US-08-389-459A-6	Sequence 6, Appl
41	34	46.6	314	3	US-08-987-867A-6	Sequence 6, Appl
42	34	46.6	327	4	US-08-679-493A-71	Sequence 71, Appl
43	34	46.6	327	4	US-08-679-493A-72	Sequence 72, Appl
44	34	46.6	357	1	US-08-468-847B-17	Sequence 17, Appl
45	34	46.6	357	4	US-09-253-316-25	Sequence 25, Appl

ALIGNMENTS

RESULT 1  
US-08-991-862-5  
; Sequence 5, Application US/08991862  
; Patent No. 6309826  
; GENERAL INFORMATION:  
; APPLICANT: Serrero, Ginette  
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
; FILE REFERENCE: Z9996.488/P001-A  
; CURRENT APPLICATION NUMBER: US/08/991,862  
; CURRENT FILING DATE: 1998-08-17  
; EARLIER APPLICATION NUMBER: 08/863,862  
; EARLIER FILING DATE: 1997-05-23  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 5:  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: mouse granulin  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(14)  
; OTHER INFORMATION: Internal peptide of mouse GP88 used to raise the  
; OTHER INFORMATION: antisera against the GP88 used in the  
; OTHER INFORMATION: immunoaffinity step.  
US-08-991-862-5

Query Match 100.0%; Score 73; DB 4; Length 14;  
Best Local Similarity 100.0%; Pred. No. 2e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SARGTKLRKKIPR 14  
| | | | | | | | | | | | | | | |  
DB 1 SARGTKLRKKIPR 14

RESULT 2  
US-07-668-648-6  
; Sequence 6, Application US/07668648  
; Patent No. 5416192  
; GENERAL INFORMATION:  
; APPLICANT: Shoyab, Mohammed  
; APPLICANT: Plozman, Gregory D.  
; TITLE OF INVENTION: NOVEL CYSTEINE-RICH GROWTH  
; TITLE OF INVENTION: MODULATING PROTEINS  
; NUMBER OF SEQUENCE ADDRESSES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York



COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/668,648  
FILING DATE: 19910819  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELEPHONE: (212)790-9090  
TELEFAX: (212) 869-9741  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 589 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-668-648-6

Query Match 100.0%; Score 73; DB 1; Length 589;  
Best Local Similarity 100.0%; Pred. No. 6.1e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SARGTKCLRKIPR 14  
Db 562 SARGTKCLRKIPR 575

RESULT 3  
US-08-429-998-6  
Sequence 6, Application US/08429998  
Patent No. 5885961  
GENERAL INFORMATION:  
APPLICANT: Shoyab, Mohammed  
APPLICANT: Plovman, Gregory D.  
TITLE OF INVENTION: EPIITHELINS: NOVEL CYSTEINE-RICH GROWTH  
TITLE OF INVENTION: MODULATING PROTEINS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/429,998  
FILING DATE: 27-APR-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/668,648  
FILING DATE: 13-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELEPHONE: (212)790-9090  
TELEFAX: (212) 869-9741  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:

LENGTH: 589 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-429-998-6

Query Match 100.0%; Score 73; DB 2; Length 589;  
Best Local Similarity 100.0%; Pred. No. 6.1e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SARGTKCLRKIPR 14  
Db 562 SARGTKCLRKIPR 575

RESULT 4  
US-08-431-333-6  
Sequence 6, Application US/08431333  
Patent No. 5965723  
GENERAL INFORMATION:  
APPLICANT: Shoyab, Mohammed  
APPLICANT: Plovman, Gregory D.  
TITLE OF INVENTION: EPIITHELINS: NOVEL CYSTEINE-RICH GROWTH  
TITLE OF INVENTION: MODULATING PROTEINS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/431,333  
FILING DATE: 27-APR-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/668,648  
FILING DATE: 13-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELEPHONE: (212)790-9090  
TELEFAX: (212) 869-9741  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 589 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-431-333-6

Query Match 100.0%; Score 73; DB 2; Length 589;  
Best Local Similarity 100.0%; Pred. No. 6.1e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SARGTKCLRKIPR 14  
Db 562 SARGTKCLRKIPR 575

RESULT 5  
US-08-991-862-2  
Sequence 2, Application US/08991862  
Patent No. 6309826  
GENERAL INFORMATION:

```
;; APPLICANT: Serrero, Ginette
;; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
;; FILE REFERENCE: Z9996.488/P001-A
;; CURRENT APPLICATION NUMBER: US/08/991,862
;; EARLIER FILING DATE: 1998-08-17
;; EARLIER APPLICATION NUMBER: 08/863,862
;; EARLIER FILING DATE: 1997-05-23
;; NUMBER OF SEQ ID NOS: 17
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 2
;; LENGTH: 589
;; TYPE: PRT
;; ORGANISM: Mouse epithelin/granulin
US-08-991-862-2

Query Match      100.0%; Score 73; DB 4; Length 589;
Best Local Similarity 100.0%; Pred. No. 6.1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SARGTKCLRKIPR 14
        |||||
DB      562 SARGTKCLRKIPR 575

RESULT 6
PCT-US91-02321-6
; Sequence 6, Application PC/TUS9102321
; GENERAL INFORMATION:
; APPLICANT: Shoyab, Mohammed
; APPLICANT: Plozman, Gregory D.
; TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH
; TITLE OF INVENTION: MODULATING PROTEINS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bristol-Myers Squibb Company
; STREET: 3005 First Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/02321
; FILING DATE: 19910403
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Poor, Brian W.
; REGISTRATION NUMBER: 32,928
; REFERENCE/DOCKET NUMBER: ON0071A-PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)728-4800
; TELEFAX: (206)448-4775
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 589 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US91-02321-6

Query Match      100.0%; Score 73; DB 5; Length 589;
Best Local Similarity 100.0%; Pred. No. 6.1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SARGTKCLRKIPR 14
        |||||
DB      562 SARGTKCLRKIPR 575
```

```
RESULT 7
US-07-668-648-2
; Sequence 2, Application US/07668648
; Patent No. 5416192
; GENERAL INFORMATION:
; APPLICANT: Shoyab, Mohammed
; APPLICANT: Plozman, Gregory D.
; TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH
; TITLE OF INVENTION: MODULATING PROTEINS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/668,648
; FILING DATE: 19910819
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 5624-161-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)790-9090
; TELEFAX: (212) 869-9741
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 589 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-668-648-2

Query Match      89.0%; Score 65; DB 1; Length 589;
Best Local Similarity 85.7%; Pred. No. 0.0014;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 SARGTKCLRKIPR 14
        |||||
DB      562 SARGTKCLRKIPR 575

RESULT 8
US-08-429-998-2
; Sequence 2, Application US/08429998
; Patent No. 5885961
; GENERAL INFORMATION:
; APPLICANT: Shoyab, Mohammed
; APPLICANT: Plozman, Gregory D.
; TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH
; TITLE OF INVENTION: MODULATING PROTEINS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```

APPLICATION NUMBER: US/08/429,998  
FILING DATE: 27-APR-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION NUMBER: US 07/668,648  
FILING DATE: 13-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELEPHONE: (212)790-9090  
TELEFAX: (212) 869-9741  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 589 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-429-998-2

Query Match 89.0%; Score 65; DB 2; Length 589;  
Best Local Similarity 85.7%; Pred. No. 0.0014;  
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SARGTKCLRRKKTTPR 14  
DB 562 SAKGTKCLRRKKTTPR 575

RESULT 9  
US-08-431-333-2  
Sequence 2, Application US/08/431333  
Patent No. 5965723  
GENERAL INFORMATION:  
APPLICANT: Shoyab, Mohammed  
APPLICANT: Plozman, Gregory D.  
TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/431,333  
FILING DATE: 27-APR-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/668,648  
FILING DATE: 13-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELEPHONE: (212)790-9090  
TELEFAX: (212) 869-9741  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 589 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-431-333-2

Query Match 89.0%; Score 65; DB 2; Length 589;  
Best Local Similarity 85.7%; Pred. No. 0.0014;  
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SARGTKCLRRKKTTPR 14  
DB 562 SAKGTKCLRRKKTTPR 575

RESULT 10  
PCT-US91-02321-2  
Sequence 2, Application PC/TUS9102321  
GENERAL INFORMATION:  
APPLICANT: Shoyab, Mohammed  
APPLICANT: Plozman, Gregory D.  
TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Bristol-Myers Squibb Company  
STREET: 3005 First Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98121  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/02321  
FILING DATE: 19910403  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Poor, Brian W.  
REGISTRATION NUMBER: 32,928  
REFERENCE/DOCKET NUMBER: ON0071A-PC  
TELEPHONE: (206)728-4800  
TELEFAX: (206)448-4775  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 589 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US91-02321-2

Query Match 89.0%; Score 65; DB 5; Length 589;  
Best Local Similarity 85.7%; Pred. No. 0.0014;  
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SARGTKCLRRKKTTPR 14  
DB 562 SAKGTKCLRRKKTTPR 575

RESULT 11  
US-07-668-648-4  
Sequence 4, Application US/07668648  
Patent No. 5416192  
GENERAL INFORMATION:  
APPLICANT: Shoyab, Mohammed  
APPLICANT: Plozman, Gregory D.  
TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York

STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/668,648  
FILING DATE: 19910819  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Misrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)790-9090  
TELEFAX: (212) 869-9741  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 593 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-668-648-4

Query Match 79.5%; Score 58; DB 1; Length 593;  
Best Local Similarity 71.4%; Pred. No. 0.023;  
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SARGTKCLRKIPR 14  
DB 566 AARGTKCLRREAPR 579

RESULT 12  
US-08-429-998-4  
Sequence 4, Application US/08429998  
Patent No. 5885961  
GENERAL INFORMATION:  
APPLICANT: Shoyab, Mohammed  
APPLICANT: Shoyab, Gregory D.  
TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH  
TITLE OF INVENTION: MODULATING PROTEINS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/429,998  
FILING DATE: 27-APR-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/668,648  
FILING DATE: 13-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Misrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)790-9090  
TELEFAX: (212) 869-9741  
INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:  
LENGTH: 593 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-429-998-4

Query Match 79.5%; Score 58; DB 2; Length 593;  
Best Local Similarity 71.4%; Pred. No. 0.023;  
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SARGTKCLRKIPR 14  
DB 566 AARGTKCLRREAPR 579

RESULT 13  
US-08-431-333-4  
Sequence 4, Application US/08431333  
Patent No. 5965793  
GENERAL INFORMATION:  
APPLICANT: Shoyab, Mohammed  
APPLICANT: Plozman, Gregory D.  
TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH  
TITLE OF INVENTION: MODULATING PROTEINS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/431,333  
FILING DATE: 27-APR-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/668,648  
FILING DATE: 13-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Misrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)790-9090  
TELEFAX: (212) 869-9741  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 593 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-431-333-4

Query Match 79.5%; Score 58; DB 2; Length 593;  
Best Local Similarity 71.4%; Pred. No. 0.023;  
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SARGTKCLRKIPR 14  
DB 566 AARGTKCLRREAPR 579

RESULT 14  
PCT-US91-02321-4  
Sequence 4, Application PC/TUS9102321  
GENERAL INFORMATION:

APPLICANT: Shoyab, Mohammed  
APPLICANT: Flowman, Gregory D.  
TITLE OF INVENTION: EPITHELIALS: NOVEL CYSTEINE-RICH GROWTH  
TITLE OF INVENTION: MODULATING PROTEINS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Bristol-Myers Squibb Company  
STREET: 3005 First Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98121  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/02321  
FILING DATE: 19910403  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Poor, Brian W.  
REGISTRATION NUMBER: 32,928  
REFERENCE/DOCKET NUMBER: ON0071A-PC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)728-4800  
TELEFAX: (206)448-4775  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 593 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US91-02321-4

Query Match 79.5%; Score 58; DB 5; Length 593;  
Best Local Similarity 71.4%; Pred. No. 0.023;  
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SARTKCLRKIPR 14  
Db 566 AARTKCLREAPR 579

## RESULT 15

US-08-991-862-7  
Sequence 7, Application US/08991862  
Patent No. 6309826  
GENERAL INFORMATION:  
APPLICANT: Sertero, Ginette  
TITLE OF INVENTION: 86 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
FILE REFERENCE: 29996.488/P001-A  
CURRENT APPLICATION NUMBER: US/08/991.862  
CURRENT FILING DATE: 1998-08-17  
EARLIER APPLICATION NUMBER: 08/863.862  
EARLIER FILING DATE: 1997-05-23  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 7  
LENGTH: 14  
TYPE: PRT  
ORGANISM: Human granulin  
FEATURE:  
NAME/KEY: PEPTIDE  
LOCATION: (1)..(14)  
OTHER INFORMATION: Internal peptide of human GP88 used to develop  
OTHER INFORMATION: neutralizing anti-human GP88 monoclonal antibody.  
US-08-991-862-7

Query Match 72.6%; Score 53; DB 4; Length 14;  
Best Local Similarity 75.0%; Pred. No. 0.0055;  
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 RGTKCLRKIPR 14  
Db 3 RGTKCLREAPR 14

Search completed: July 8, 2003, 16:30:19  
Job time : 9.61539 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2003, 16:24:26 ; Search time 13.641 Seconds

(without alignments)  
119.483 Million cell updates/sec

Title: US-09-824-647-5

Perfect score: 73

Sequence: 1 SARGTKCLRKIKPR 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 116419773 residues

Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB\_PEP.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB\_PEP.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB\_PEP.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB\_PEP.\*  
5: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB\_PEP.\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB\_PEP.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB\_PEP.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB\_PEP.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB\_PEP.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB\_PEP.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB\_PEP.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB\_PEP.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB\_PEP.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB\_PEP.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	73	100.0	14	9	US-09-824-647-5
2	73	100.0	14	9	US-10-218-509-5
3	73	100.0	14	9	US-10-281-160-5
4	73	100.0	14	10	US-09-813-156-5
5	73	100.0	14	10	US-09-824-807-5
6	73	100.0	589	9	US-09-824-647-2
7	73	100.0	589	9	US-10-218-509-2
8	73	100.0	589	9	US-10-281-160-2
9	73	100.0	589	10	US-09-813-156-2
10	73	100.0	589	10	US-09-824-807-2
11	58	79.5	621	10	US-09-925-301-1416
12	53	72.6	14	9	US-09-824-647-7
13	53	72.6	14	9	US-10-218-509-7
14	53	72.6	14	9	US-10-281-160-7
15	53	72.6	14	10	US-09-813-156-7
16	53	72.6	14	10	US-09-824-807-7
17	53	72.6	593	9	US-09-824-647-17
18	53	72.6	593	9	US-10-218-509-17
19	53	72.6	593	9	US-10-281-160-17

20	53	72.6	593	10	US-09-813-156-17	Sequence 17, Appl
21	53	72.6	593	10	US-09-824-807-17	Sequence 17, Appl
22	40	54.8	255	9	US-09-866-050A-692	Sequence 692, App
23	40	54.8	275	9	US-10-269-353-15	Sequence 15, Appl
24	40	54.8	275	10	US-09-790-264-15	Sequence 15, Appl
25	39	53.4	41	10	US-09-872-852-3	Sequence 3, Appl
26	39	53.4	45	10	US-09-872-852-4	Sequence 4, Appl
27	39	53.4	67	10	US-09-917-340-52	Sequence 52, Appl
28	39	53.4	67	10	US-09-917-340-72	Sequence 72, Appl
29	39	53.4	67	10	US-09-872-852-2	Sequence 2, Appl
30	39	53.4	76	9	US-09-738-626-5670	Sequence 5670, Ap
31	38	52.1	263	10	US-09-840-704-5	Sequence 5, Appl
32	37	50.7	106	9	US-10-091-438-179	Sequence 179, App
33	37	50.7	106	10	US-09-764-853-697	Sequence 697, App
34	36	49.3	51	9	US-09-764-891-4995	Sequence 4995, Ap
35	36	49.3	407	9	US-09-769-787-14	Sequence 14, Appl
36	36	49.3	520	9	US-10-157-855-2	Sequence 2, Appl
37	36	49.3	520	9	US-10-223-076-3	Sequence 3, Appl
38	36	49.3	598	9	US-09-900-797-26	Sequence 26, Appl
39	36	49.3	598	10	US-09-887-586A-26	Sequence 26, Appl
40	36	49.3	598	10	US-09-903-012-26	Sequence 26, Appl
41	35	47.9	41	9	US-10-277-693A-7	Sequence 7, Appl
42	35	47.9	51	10	US-09-864-761-41212	Sequence 41212, A
43	35	47.9	148	9	US-10-226-489-11	Sequence 11, Appl
44	35	47.9	550	10	US-09-934-323-6	Sequence 6, Appl
45	35	47.9	835	10	US-09-934-323-2	Sequence 2, Appl

#### ALIGNMENTS

RESULT 1  
US-09-824-647-5  
; Sequence 5, Application US/09824647  
; Publication No. US20020183270A1  
; GENERAL INFORMATION:  
; APPLICANT: Serrero, Ginette  
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
; FILE REFERENCE: Z9996.488/P001-A  
; CURRENT APPLICATION NUMBER: US/09/824,647  
; CURRENT FILING DATE: 2001-04-04  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/991,862  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/863,862  
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-05-23  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: mouse granuln  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(14)  
; OTHER INFORMATION: Internal peptide of mouse GP88 used to raise the  
; OTHER INFORMATION: antiserum against the GP88 used in the  
; OTHER INFORMATION: immunoaffinity step.  
US-09-824-647-5

Query Match 100.0%; Score 73; DB 9; Length 14;  
Best Local Similarity 100.0%; Pred. No. 1.9e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SARGTKCLRKIKPR 14  
| | | | | | | | | | | | | | | |  
Db 1 SARGTKCLRKIKPR 14

RESULT 2  
US-10-218-509-5  
; Sequence 5, Application US/10218509  
; Publication No. US20030092661A1  
; GENERAL INFORMATION:

; APPLICANT: Seirero, Ginette  
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
; FILE REFERENCE: 29996.488/P001-A  
; CURRENT APPLICATION NUMBER: US/10/218,509  
; CURRENT FILING DATE: 2002-08-15  
; PRIOR APPLICATION NUMBER: 08/991,862  
; PRIOR FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: 08/863,862  
; PRIOR FILING DATE: 1997-05-23  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: mouse granulin  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(14)  
; OTHER INFORMATION: Internal peptide of mouse GP88 used to raise the  
; OTHER INFORMATION: antisera against the GP88 used in the  
; OTHER INFORMATION: immunoaffinity step.  
US-10-218-509-5

Query Match 100.08; Score 73; DB 9; Length 14;  
Best Local Similarity 100.08; Pred. No. 1.9e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SARGTKCLRKIKIPR 14  
| | | | | | | | | | | | | | | |  
DB 1 SARGTKCLRKIKIPR 14

## RESULT 3

US-10-281-160-5  
; Sequence 5, Application US/10281160  
; Publication No. US20030108950A1  
; GENERAL INFORMATION:  
; APPLICANT: Seirero, Ginette  
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
; FILE REFERENCE: 29996.488/P001-A  
; CURRENT APPLICATION NUMBER: US/10/281,160  
; CURRENT FILING DATE: 2002-10-28  
; PRIOR APPLICATION NUMBER: 08/991,862  
; PRIOR FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: 08/863,862  
; PRIOR FILING DATE: 1997-05-23  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: mouse granulin  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(14)  
; OTHER INFORMATION: Internal peptide of mouse GP88 used to raise the  
; OTHER INFORMATION: antisera against the GP88 used in the  
; OTHER INFORMATION: immunoaffinity step.  
US-10-281-160-5

Query Match 100.08; Score 73; DB 9; Length 14;  
Best Local Similarity 100.08; Pred. No. 1.9e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SARGTKCLRKIKIPR 14  
| | | | | | | | | | | | | | | |  
DB 1 SARGTKCLRKIKIPR 14

## RESULT 4

US-09-813-156-5  
; Sequence 5, Application US/09813156  
; Patent No. US20020061859A1

; GENERAL INFORMATION:  
; APPLICANT: Seirero, Ginette  
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
; FILE REFERENCE: 29996.488/P001-A  
; CURRENT APPLICATION NUMBER: US/09/813,156  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 08/991,862  
; PRIOR FILING DATE: 1997-12-16  
; PRIOR APPLICATION NUMBER: 08/863,862  
; PRIOR FILING DATE: 1997-05-23  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: mouse granulin  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(14)  
; OTHER INFORMATION: Internal peptide of mouse GP88 used to raise the  
; OTHER INFORMATION: antisera against the GP88 used in the  
; OTHER INFORMATION: immunoaffinity step.  
US-09-813-156-5

Query Match 100.08; Score 73; DB 10; Length 14;  
Best Local Similarity 100.08; Pred. No. 1.9e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SARGTKCLRKIKIPR 14  
| | | | | | | | | | | | | | | |  
DB 1 SARGTKCLRKIKIPR 14

## RESULT 5

US-09-824-807-5  
; Sequence 5, Application US/09824807  
; Patent No. US20020094966A1  
; GENERAL INFORMATION:  
; APPLICANT: Seirero, Ginette  
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
; FILE REFERENCE: 29996.488/P001-A  
; CURRENT APPLICATION NUMBER: US/09/824,807  
; CURRENT FILING DATE: 2001-04-04  
; PRIOR APPLICATION NUMBER: 08/991,862  
; PRIOR FILING DATE: 1997-12-16  
; PRIOR APPLICATION NUMBER: 08/863,862  
; PRIOR FILING DATE: 1997-05-23  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: mouse granulin  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(14)  
; OTHER INFORMATION: Internal peptide of mouse GP88 used to raise the  
; OTHER INFORMATION: antisera against the GP88 used in the  
; OTHER INFORMATION: immunoaffinity step.  
US-09-824-807-5

Query Match 100.08; Score 73; DB 10; Length 14;  
Best Local Similarity 100.08; Pred. No. 1.9e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SARGTKCLRKIKIPR 14  
| | | | | | | | | | | | | | | |  
DB 1 SARGTKCLRKIKIPR 14

## RESULT 6

US-09-824-647-2  
; Sequence 2, Application US/09824647

```
; Publication No. US20020183270A1
; GENERAL INFORMATION:
; APPLICANT: Serrero, Ginette
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
; FILE REFERENCE: 29996.488/P001-A
; CURRENT APPLICATION NUMBER: US/09/824,647
; CURRENT FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/991,862
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/863,862
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-05-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Mouse epithelin/granulin
US-09-824-647-2

Query Match      100.0%; Score 73; DB 9; Length 589;
Best Local Similarity 100.0%; Pred. No. 9e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SARGTKCLRRKKIPR 14
DB      562 SARGTKCLRRKKIPR 575

RESULT 7
US-10-218-509-2
; Sequence 2, Application US/10218509
; Publication No. US20030092661A1
; GENERAL INFORMATION:
; APPLICANT: Serrero, Ginette
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
; FILE REFERENCE: 29996.488/P001-A
; CURRENT APPLICATION NUMBER: US/10/218,509
; CURRENT FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: 08/991,862
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 08/863,862
; PRIOR FILING DATE: 1997-05-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Mouse epithelin/granulin
US-10-218-509-2

Query Match      100.0%; Score 73; DB 9; Length 589;
Best Local Similarity 100.0%; Pred. No. 9e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SARGTKCLRRKKIPR 14
DB      562 SARGTKCLRRKKIPR 575

RESULT 8
US-10-281-160-2
; Sequence 2, Application US/10281160
; Publication No. US20030108950A1
; GENERAL INFORMATION:
; APPLICANT: Serrero, Ginette
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
; FILE REFERENCE: 29996.488/P001-A
; CURRENT APPLICATION NUMBER: US/10/281,160
; CURRENT FILING DATE: 2002-10-28
; PRIOR APPLICATION NUMBER: US/08/991,862
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 08/863,862
; PRIOR FILING DATE: 1997-05-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Mouse epithelin/granulin
US-09-824-807-2

Query Match      100.0%; Score 73; DB 10; Length 589;
Best Local Similarity 100.0%; Pred. No. 9e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Mouse epithelin/granulin
US-10-281-160-2

Query Match      100.0%; Score 73; DB 9; Length 589;
Best Local Similarity 100.0%; Pred. No. 9e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SARGTKCLRRKKIPR 14
DB      562 SARGTKCLRRKKIPR 575

RESULT 9
US-09-813-156-2
; Sequence 2, Application US/09813156
; Patent No. US20020061859A1
; GENERAL INFORMATION:
; APPLICANT: Serrero, Ginette
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
; FILE REFERENCE: 29996.488/P001-A
; CURRENT APPLICATION NUMBER: US/09/813,156
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 08/991,862
; PRIOR FILING DATE: 1997-12-16
; PRIOR APPLICATION NUMBER: 08/863,862
; PRIOR FILING DATE: 1997-05-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Mouse epithelin/granulin
US-09-813-156-2

Query Match      100.0%; Score 73; DB 10; Length 589;
Best Local Similarity 100.0%; Pred. No. 9e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SARGTKCLRRKKIPR 14
DB      562 SARGTKCLRRKKIPR 575

RESULT 10
US-09-824-807-2
; Sequence 2, Application US/09824807
; Patent No. US20020094966A1
; GENERAL INFORMATION:
; APPLICANT: Serrero, Ginette
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
; FILE REFERENCE: 29996.488/P001-A
; CURRENT APPLICATION NUMBER: US/09/824,807
; CURRENT FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 08/991,862
; PRIOR FILING DATE: 1997-12-16
; PRIOR APPLICATION NUMBER: 08/863,862
; PRIOR FILING DATE: 1997-05-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Mouse epithelin/granulin
US-09-824-807-2

Query Match      100.0%; Score 73; DB 10; Length 589;
Best Local Similarity 100.0%; Pred. No. 9e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



QY 1 SARGTKCLRRKKIPR 14  
Db 562 SARGTKCLRRKKIPR 575

RESULT 11  
US-09-925-301-1416  
; Sequence 1416, Application US/09925301  
; Patent No. US20020052308A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: P106  
; CURRENT APPLICATION NUMBER: US/09/925,301  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05882  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1694  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 1416  
; LENGTH: 621  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-925-301-1416

Query Match 79.5%; Score 58; DB 10; Length 621;  
Best Local Similarity 71.4%; Pred. No. 0.043;  
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SARGTKCLRRKKIPR 14  
Db 594 AARGTKCLRRREAPR 607

RESULT 12  
US-09-824-647-7  
; Sequence 7, Application US/09824647  
; Publication No. US20020183270A1  
; GENERAL INFORMATION:  
; APPLICANT: Seriero, Ginette  
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
; FILE REFERENCE: 29996.488/P001-A  
; CURRENT APPLICATION NUMBER: US/09/824,647  
; CURRENT FILING DATE: 2001-04-04  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/991,862  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/863,862  
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-05-23  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Human granulin  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(14)  
; OTHER INFORMATION: Internal peptide of human GP88 used to develop  
; OTHER INFORMATION: neutralizing anti-human GP88 monoclonal antibody.  
US-09-824-647-7

Query Match 72.6%; Score 53; DB 9; Length 14;  
Best Local Similarity 75.0%; Pred. No. 0.0066;  
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RGTKCLRRKKIPR 14  
Db 3 RGTKCLRRREAPR 14

RESULT 13  
US-10-218-509-7  
; Sequence 7, Application US/10218509  
; Publication No. US20030092661A1  
; GENERAL INFORMATION:  
; APPLICANT: Seriero, Ginette  
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
; FILE REFERENCE: 29996.488/P001-A  
; CURRENT APPLICATION NUMBER: US/10/218,509  
; CURRENT FILING DATE: 2002-08-15  
; PRIOR APPLICATION NUMBER: 08/991,862  
; PRIOR FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: 08/863,862  
; PRIOR FILING DATE: 1997-05-23  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Human granulin  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(14)  
; OTHER INFORMATION: Internal peptide of human GP88 used to develop  
; OTHER INFORMATION: neutralizing anti-human GP88 monoclonal antibody.  
US-10-218-509-7

Query Match 72.6%; Score 53; DB 9; Length 14;  
Best Local Similarity 75.0%; Pred. No. 0.0066;  
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RGTKCLRRKKIPR 14  
Db 3 RGTKCLRRREAPR 14

RESULT 14  
US-10-281-160-7  
; Sequence 7, Application US/10281160  
; Publication No. US20030108950A1  
; GENERAL INFORMATION:  
; APPLICANT: Seriero, Ginette  
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
; FILE REFERENCE: 29996.488/P001-A  
; CURRENT APPLICATION NUMBER: US/10/281,160  
; CURRENT FILING DATE: 2002-10-28  
; PRIOR APPLICATION NUMBER: 08/991,862  
; PRIOR FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: 08/863,862  
; PRIOR FILING DATE: 1997-05-23  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Human granulin  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(14)  
; OTHER INFORMATION: Internal peptide of human GP88 used to develop  
; OTHER INFORMATION: neutralizing anti-human GP88 monoclonal antibody.  
US-10-281-160-7

Query Match 72.6%; Score 53; DB 9; Length 14;  
Best Local Similarity 75.0%; Pred. No. 0.0066;  
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RGTKCLRRKKIPR 14  
Db 3 RGTKCLRRREAPR 14

RESULT 15

US-09-813-156-7  
; Sequence 7, Application US/09813156  
; Patent No. US20020061859A1  
; GENERAL INFORMATION:  
; APPLICANT: Seriero, Ginette  
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
; FILE REFERENCE: 29996.488/P001-A  
; CURRENT APPLICATION NUMBER: US/09/813,156  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 08/991,862  
; PRIOR FILING DATE: 1997-12-16  
; PRIOR APPLICATION NUMBER: 08/863,862  
; PRIOR FILING DATE: 1997-05-23  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Human granulin  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(14)  
; OTHER INFORMATION: Internal peptide of human GP88 used to develop  
; OTHER INFORMATION: neutralizing anti-human GP88 monoclonal antibody.  
US-09-813-156-7

Query Match 72.6%; Score 53; DB 10; Length 14;  
Best Local Similarity 75.0%; Pred. No. 0.0056;  
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RGTKCLRKIKPR 14  
| | | | | | | | | |  
DB 3 RGTKCLRREAPR 14

Search completed: July 8, 2003, 16:31:41  
Job time: 14.641 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2003, 16:19:45 ; Search time 14.1795 Seconds  
(without alignments)  
94.918 Million cell updates/sec

Title: US-09-824-647-5

Perfect score: 73  
Sequence: 1 SARGTKCLRKIKIPR 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_73:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	73	100.0	589	2 C38128	epithelin/granulin
2	65	89.0	589	2 B38128	epithelin/granulin
3	58	79.5	593	1 GYHU	granulin precursor
4	41	56.2	452	2 G64476	hypothetical prote
5	40	54.8	289	2 JC7279	Down syndrome crit
6	40	54.8	555	2 AB0321	glutamine-tRNA lig
7	39	53.4	288	2 T24824	hypothetical prote
8	39	53.4	424	2 E64450	hypothetical prote
9	39	53.4	942	2 S75598	poly(A) polymerase
10	38	52.1	411	2 AH0361	probable aminotran
11	38	52.1	821	2 T48400	serine/threonine-p
12	38	52.1	825	2 T02518	hypothetical prote
13	37.5	51.4	341	2 C86246	hypothetical prote
14	37	50.7	283	2 C71554	probable polysacch
15	37	50.7	327	2 T49514	hypothetical prote
16	37	50.7	331	2 T28023	hypothetical prote
17	37	50.7	422	2 T49513	gastric mucin rela
18	37	50.7	542	2 AE3057	phosphoglucosidase
19	37	50.7	550	2 D64354	queine tRNA-ribos
20	37	50.7	566	2 AH3487	phosphoglucosidase
21	37	50.7	567	2 A96229	phosphoglucosidase
22	36.5	50.0	533	2 T19416	hypothetical prote
23	36	49.3	66	2 F97737	hypothetical prote
24	36	49.3	158	2 B82299	transcription elon
25	36	49.3	231	2 T01294	sterol O-acyltrans
26	36	49.3	265	2 S60947	hypothetical prote
27	36	49.3	336	2 T50560	SINAH1 protein [m
28	36	49.3	422	2 B82272	integrase, phage f
29	36	49.3	441	2 H84576	diacylglycerol O-a

## ALIGNMENTS

### RESULT 1

C38128

epithelin/granulin precursor - mouse  
N:Alternate names: acrogranin; PC-cell-derived growth factor  
C:Species: Mus musculus (house mouse)  
C>Date: 10-Jul-1992 #sequence\_revision 10-Jul-1992 #text\_change 20-Aug-1999  
C:Accession: C38128; S32503; I49468; A46705  
R:Plowman, G.D.; Green, J.M.; Neubauer, M.G.; Buckley, S.D.; McDonald, V.L.; Toda  
J. Biol. Chem. 267, 13073-13078, 1992  
A:Title: The epithelin precursor encodes two proteins with opposing activities on  
A:Reference number: A38128; MUID:92317004; PMID:1618805  
A:Accession: C38128  
A:Molecule type: mRNA  
A:Residues: 1-589 <PIQ>  
A:Cross-references: GB:X62321; NID:G50851; PIDN:CAA44197.1; PID:G50852  
R:Baba, T.; Nemoto, H.; Watanabe, K.; Arai, Y.; Gerton, G.L.  
FEBS Lett. 322, 89-94, 1993  
A:Title: Exon/Intron organization of the gene encoding the mouse epithelin/granul  
A:Reference number: S32503; MUID:93245991; PMID:8482392  
A:Accession: S32503  
A:Molecule type: DNA  
A:Residues: 18-349, 'L', 351-589 <BAE>  
R:Baba, T.; Hoff, H.B.  
Mol. Reprod. Dev. 34, 233-243, 1993  
A:Title: Acrogranin, an acrosomal cysteine-rich glycoprotein, is the precursor of  
A:Reference number: I48141; MUID:93228994; PMID:8471244  
A:Accession: I49468  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-250, 'L', 252-253, 'V', 255-349, 'L', 351-401, 'SA', 404-589 <RES>  
A:Cross-references: GB:M66736; NID:G191766; PIDN:AAA37191.1; PID:G191767  
R:Zhou, J.; Gao, G.; Crabb, J.W.; Serrero, G.  
J. Biol. Chem. 268, 10863-10869, 1993  
A:Title: Purification of an autocrine growth factor homologous with mouse epithel  
A:Reference number: A46705; MUID:93266526; PMID:8496151  
A:Accession: A46705  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 18-19, 'X', 21-25, 'X', 27-29, 'XX', 32, 'XXX', 119-127; 152-154, 'DXK', 158-161  
C:Superfamily: granulin

Query Match 100.0%; Score 73; DB 2; Length 589;  
Best Local Similarity 100.0%; Pred. No. 2e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SARGTKCLRKIKIPR 14  
|||||

Db 562 SARGTKCLRKIKIPR 575

### RESULT 2

B38128

epithelin/granulin precursor - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 10-Jul-1992 #sequence\_revision 10-Jul-1992 #text\_change 20-Aug-1999  
 C:Accession: B38128; A36199; B36199; E36698; I53272  
 R:Plowman, G.D.; Green, J.M.; Neubauer, M.G.; Buckley, S.D.; McDonald, V.L.; Todaro, G.J.  
 J. Biol. Chem. 267, 13073-13078, 1992  
 A:Title: The epithelin precursor encodes two proteins with opposing activities on epithelial cells  
 A:Reference number: A38128; MUID:92317004; PMID:1618805  
 A:Accession: B38128  
 A:Molecule type: mRNA  
 A:Residues: 1-589 <P>  
 A:Cross-references: GB:X62322; MID:956108; PIDN:CAA44198.1; PID:956109  
 R:Shoyab, M.; McDonald, V.L.; Byles, C.; Todaro, G.J.; Plowman, G.D.  
 Proc. Natl. Acad. Sci. U.S.A. 87, 7912-7916, 1990  
 A:Title: Epithelins 1 and 2: isolation and characterization of two cysteine-rich growth factors  
 A:Reference number: A36199; MUID:91045907; PMID:2236009  
 A:Accession: A36199  
 A:Molecule type: protein  
 A:Residues: 280-300 <SHO>  
 A:Accession: B36199  
 A:Molecule type: protein  
 A:Residues: 205-226 <SH2>  
 R:Batemán, A.; Belcourt, D.; Bennett, H.; Lazure, C.; Solomon, S.  
 Biochem. Biophys. Res. Commun. 173, 1161-1168, 1990  
 A:Title: Granulins, a novel class of peptide from leukocytes.  
 A:Reference number: A36698; MUID:91097544; PMID:2268320  
 A:Accession: E36698  
 A:Molecule type: protein  
 A:Residues: 279-307, SB', 310-324, 'T', 326, 'X', 328, 'Q' <BAT>  
 R:Bhandari, V.; Gaid, A.; Bateman, A.  
 Endocrinology 133, 2682-2689, 1993  
 A:Title: The complementary deoxyribonucleic acid sequence, tissue distribution, and cell expression of the human granulins  
 A:Reference number: I53272; MUID:94062640; PMID:8243292  
 A:Accession: I53272  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-200, 'S', 203-388, 'M', 390-589 <RES>  
 A:Cross-references: GB:M97750; MID:9204223; PIDN:AAA16903.1; PID:g204224  
 C:Superfamily: granulin

Query Match 89.0%; Score 65; DB 2: Length 589;  
 Best Local Similarity 85.7%; Pred. No. 0.00058;  
 Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SARGTKCLRKIPR 14  
 II:|||||II  
 Db 562 SAKGTKCLRKIPR 575

## RESULT 3

GHU  
 granulin precursor [validated] - human  
 N:Alternate names: epithelin  
 N:Contains: granulin A; granulin B; granulin C; granulin D; granulin E; granulin F; granulin G  
 C:Species: Homo sapiens (man)  
 C:Date: 30-Sep-1992 #sequence\_revision 03-May-1996 #text\_change 08-Dec-2000  
 C:Accession: JCI1284; A38128; A36199; B36698; D36698; E36698; I53272  
 R:Bhandari, V.; Bateman, A.  
 Biochem. Biophys. Res. Commun. 188, 57-63, 1992  
 A:Title: Structure and chromosomal location of the human granulin gene.  
 A:Reference number: JCI1284; MUID:93038704; PMID:1417868  
 A:Accession: JCI1284  
 A:Molecule type: DNA  
 A:Residues: 1-593 <BHA>  
 R:Plowman, G.D.; Green, J.M.; Neubauer, M.G.; Buckley, S.D.; McDonald, V.L.; Todaro, G.J.  
 J. Biol. Chem. 267, 13073-13078, 1992  
 A:Title: The epithelin precursor encodes two proteins with opposing activities on epithelial cells  
 A:Reference number: A38128; MUID:92317004; PMID:1618805  
 A:Accession: A38128  
 A:Molecule type: mRNA  
 A:Residues: 1-593 <P>  
 A:Cross-references: GB:X62320; MID:931192; PIDN:CAA44196.1; PID:g31193

R:Bhandari, V.; Palfrey, R.G.E.; Bateman, A.  
 Proc. Natl. Acad. Sci. U.S.A. 89, 1715-1719, 1992  
 A:Title: Isolation and sequence of the granulin precursor cDNA from human bone marrow  
 A:Reference number: A38118; MUID:92179253; PMID:1542665  
 A:Accession: A38118  
 A:Molecule type: mRNA  
 A:Residues: 1-406, 'R', 408-433, 'G', 435-453, 'G', 455-459, 'Q', 461-546, 'A', 548-566, 'R', 'S'  
 A:Cross-references: GB:M75161; MID:9183612; PIDN:AAA58617.1; PID:g183613  
 A:Note: This sequence has been revised in reference JCI1284  
 R:Batemán, A.; Belcourt, D.; Bennett, H.; Lazure, C.; Solomon, S.  
 Biochem. Biophys. Res. Commun. 173, 1161-1168, 1990  
 A:Title: Granulins, a novel class of peptide from leukocytes.  
 A:Reference number: A36698; MUID:91097544; PMID:2268320  
 A:Accession: A36698  
 A:Molecule type: protein  
 A:Residues: 281-336 <BAT>  
 A:Note: This protein was purified and characterized as granulin A  
 A:Accession: B36698  
 A:Molecule type: protein  
 A:Residues: 206-218, 'H', 220-233 <BA2>  
 A:Note: This protein was purified and characterized as granulin B  
 A:Accession: C36698  
 A:Molecule type: protein  
 A:Residues: 364-367, 'X', 369-385, 'H', 387-396 <BA3>  
 A:Note: This protein was purified and characterized as granulin C  
 A:Accession: D36698  
 A:Molecule type: protein  
 A:Residues: 442-445, 'XDTSS', 456-458, 'DG', <BA4>  
 R:Kardana, A.; Bagshawe, K.D.; Coles, B.; Read, D.; Taylor, M.  
 Br. J. Cancer 67, 686-692, 1993  
 A:Title: Characterisation of UGP and its relationship with beta-core fragment.  
 A:Reference number: A56873; MUID:93229246; PMID:8471426  
 A:Accession: A56873  
 A:Molecule type: protein  
 A:Residues: 281-283, 'X', 285-289, 'S', 291-295 <KAR>  
 A:Experimental source: urine  
 A:Note: sequence extracted from NCBI backbone (NCBIP:129524)  
 C:Genetics:  
 A:Gene: GDB:GRN  
 A:Cross-references: GDB:136006; OMIM:138945  
 A:Map position: 17pter-17qter  
 A:Introns: 46/3; 88/3; 117/1; 154/3; 200/1; 236/3; 279/1; 311/3; 393/3; 471/3; 548/3; 548/3  
 C:Superfamily: granulin  
 C:Keywords: glycoprotein; tandem repeat  
 F:1-17/Domain: signal sequence #status predicted <SIG>  
 F:18-593/Product: granulin #status predicted <MAP>  
 F:18-593/Product: programulin #status predicted <PRO>  
 F:18-44/Product: paraganulin #status experimental <PCR>  
 F:58-113/Product: granulin G #status predicted <GRG>  
 F:123-179/Product: granulin F #status predicted <GRF>  
 F:206-261/Product: granulin B #status experimental <GRB>  
 F:281-336/Product: granulin A #status experimental <GRA>  
 F:364-417/Product: granulin C #status experimental <GRC>  
 F:442-496/Product: granulin D #status predicted <GRD>  
 F:518-573/Product: granulin E #status predicted <GRE>  
 F:368/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 79.5%; Score 58; DB 1: Length 593;  
 Best Local Similarity 71.4%; Pred. No. 0.011;  
 Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 SARGTKCLRKIPR 14  
 :|||||II  
 Db 566 AARGTKCLRKIPR 579

## RESULT 4

G64476  
 hypothetical protein MJ1416 - Methanococcus jannaschii  
 C:Species: Methanococcus jannaschii  
 C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 29-Sep-1999  
 C:Accession: G64476  
 R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Bl

; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstein, K.G.; Merrick, J.M.; Glodek, A.;  
 rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.

Science 273, 1058-1073, 1996  
 A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.  
 A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii

A:Reference number: A64300; MUID:96337999; PMID:8688087

A:Accession: G64476

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-452 <BUIL>

A:Cross-references: GB:067582; GB:L77117; NID:gl592064; PIDN:AB99426.1; PID:gl500295; T

C:Genetics:

A:Map position: REV1378352-1376994

C:Superfamily: conserved hypothetical protein MJ1416

Query Match 56.28; Score 41; DB 2; Length 452;

Best Local Similarity 87.58; Pred. No. 12;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 KCLRKKIP 13

Db 10 RCLRKKIP 17

|||||

RESULT 5

JC7279

Down syndrome critical region gene-2 (DSCR2) protein - mouse

N:Alternate names: chromosome 21-leucine rich protein (C21-LRP)

C:Species: Mus musculus (house mouse)

C:Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 02-Sep-2000

A:Accession: JC7279

R:Vidal-Taboada, J.M.; Lu, A.; Pique, M.; Pons, G.; Gil, J.; Oliva, R.

Biochem. Biophys. Res. Commun. 272, 156-163, 2000

A:Title: Down syndrome critical region gene 2: Expression during mouse development and i

A:Reference number: JC7279

A:Accession: JC7279

A:Molecule type: mRNA

A:Residues: 1-289 <VID>

A:Cross-references: GB:AJ238270

C:Genetics:

A:Gene: dscr2/c21-lrp

A:Map position: 21

C:Superfamily: human 32.9K leucine-rich protein (chromosome 21)

C:Keywords: Down's syndrome; transmembrane protein

Query Match 54.8%; Score 40; DB 2; Length 289;

Best Local Similarity 61.5%; Pred. No. 13;

Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 SARGTKCLRKKIP 13

Db 256 SSRLKCLVKNIP 268

|||||

RESULT 6

AB0321

glutamine-tRNA ligase (EC 6.1.1.18) [imported] - Yersinia pestis (strain CO92)

C:Species: Yersinia pestis

C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 03-Jun-2002

A:Accession: AB0321

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;

Il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,

Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AB0321

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-555 <KUP>

A:Cross-references: GB:AL590842; PIDN:CAC92873.1; PID:gl5980617; GSPDB:GN00175

C:Genetics:

A:Gene: glns

C:Superfamily: glutamine-tRNA ligase; glutamine-tRNA ligase homology  
 C:Keywords: ligase

Query Match 54.8%; Score 40; DB 2; Length 555;

Best Local Similarity 72.7%; Pred. No. 22;

Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 ARGTKCLRKKI 12

Db 167 AEGTACLRKI 177

|||||

RESULT 7

T24824

hypothetical protein T11A5.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

A:Accession: T24824

R:McMurray, A.

submitted to the EMBL Data Library, May 1996

A:Reference number: Z19939

A:Accession: T24824

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-288 <WIL>

A:Cross-references: EMBL:Z72515; PIDN:CAA96685.1; GSPDB:GN00023; CESP:T11A5.5

A:Experimental source: clone T11A5

C:Genetics:

A:Gene: CESP:T11A5.5

A:Map position: 5

A:Introns: 202/2

Query Match 53.4%; Score 39; DB 2; Length 288;

Best Local Similarity 70.0%; Pred. No. 19;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 GTKCLRKKIP 13

Db 145 GSKCLLRKAP 154

|||||

RESULT 8

E64450

hypothetical protein MJ1206 - Methanococcus jannaschii

C:Species: Methanococcus jannaschii

C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000

A:Accession: E64450

R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.;

Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glo

ron, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.

Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.;

A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jai

A:Reference number: A64300; MUID:96337999; PMID:8688087

A:Accession: E64450

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-424 <BUIL>

A:Cross-references: GB:U67562; GB:L77117; NID:g2826374; PIDN:AB999210.1; PID:gl51

C:Genetics:

A:Map position: REV1149948-1148674

A:Start codon: GTG

C:Superfamily: conserved hypothetical protein MJ1206

Query Match 53.4%; Score 39; DB 2; Length 424;

Best Local Similarity 87.5%; Pred. No. 27;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 KCLRKKIP 13

Db 259 KCLRSKIP 266

|||||

## RESULT 9

S75598  
 POLY(A) polymerase - *Synechocystis* sp. (strain PCC 6803)  
 N:Alternate names: protein sll1253  
 C:Species: *Synechocystis* sp.  
 A:Variety: PCC 6803  
 C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999  
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asanizu, E.; Nakamura, Y.; Miyaajima, N.; O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda  
 DNA Res. 3, 109-136, 1996  
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*  
 S.  
 A:Reference number: S74322; MUID:97061201; PMID:8905231  
 A:Accession: S75598  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-942 <KAN>  
 A:Cross-references: EMBL:D90912; GB:AB001339; NID:gi653228; PIDN:BAAL18159.1; PID:d101889  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
 C:Genetics:  
 A:Gene: pcnB

Query Match 53.4%; Score 39; DB 2; Length 942;  
 Best Local Similarity 66.7%; Pred. No. 54;  
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SARGTKCLRKXI 12  
 :||| | |||:  
 DB 865 AARGNKLKQI 876

## RESULT 10

AH0361  
 Probable aminotransferase YPO2975 [imported] - *Yersinia pestis* (strain CO92)  
 C:Species: *Yersinia pestis*  
 C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Nov-2001  
 C:Accession: AH0361  
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
 11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,  
 Nature 413, 523-527, 2001  
 A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.  
 A:Reference number: AB0001; MUID:21470413; PMID:11586360  
 A:Accession: AH0361  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-411 <KUR>  
 A:Cross-references: GB:AL590842; PIDN:CAC92219.1; PID:gl5980931; GSPDB:GN00175  
 C:Genetics:  
 A:Gene: YPO2975  
 C:Superfamily: aspartate transaminase

Query Match 52.1%; Score 38; DB 2; Length 411;  
 Best Local Similarity 50.0%; Pred. No. 40;  
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 SARGTKCLRKXIPR 14  
 :||| |||:  
 DB 72 TSGIPCLKRAISR 85

## RESULT 11

T48400  
 serine/threonine-protein kinase ctrl - *Arabidopsis thaliana*  
 N:Alternate names: protein F17C15.150  
 C:Species: *Arabidopsis thaliana* (mouse-ear cross)  
 C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 17-Nov-2000  
 C:Accession: T48400; A45178  
 R:Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.;  
 submitted to the Protein Sequence Database, March 2000  
 A:Reference number: Z24492  
 A:Accession: T48400

A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-821 <BEV>  
 A:Cross-references: EMBL:AL162506  
 A:Experimental source: cultivar Columbia; BAC clone F17C15  
 R:Kieber, J.J.; Rothenberg, M.; Roman, G.; Feldmann, K.A.; Ecker, J.R.  
 Cell 72, 427-441, 1993  
 A:Title: CTRL, a negative regulator of the ethylene response pathway in *Arabidopsis*  
 A:Reference number: A45178; MUID:93161417; PMID:8431946  
 A:Contents: Columbia  
 A:Accession: A45178  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-468; 470-821 <KIE>  
 A:Note: sequence extracted from NCBI backbone (NCBIP:124878)  
 C:Genetics:  
 A:Map position: 5  
 A:Introns: 216/3; 317/2; 364/3; 399/2; 468/1; 560/1; 573/3; 596/3; 630/2; 658/3; 69  
 A:Note: F17C15.150  
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase b  
 C:Keywords: ATP  
 F:549-812/Domain: protein kinase homology <KIN>  
 F:557-565/Region: protein kinase ATP-binding motif

Query Match 52.1%; Score 38; DB 2; Length 821;  
 Best Local Similarity 57.1%; Pred. No. 74;  
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 SARGTKCLRKXIPR 14  
 :||| | |||:  
 DB 761 AAVGFKRLEIPR 774

## RESULT 12

T02518  
 Hypothetical protein At2g37520 [imported] - *Arabidopsis thaliana*  
 N:Alternate names: hypothetical protein F13M22.2  
 C:Species: *Arabidopsis thaliana* (mouse-ear cross)  
 C:Date: 05-Mar-1999 #sequence\_revision 05-Mar-1999 #text\_change 16-Feb-2001  
 C:Accession: T02518; G84793  
 R:Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sy  
 submitted to the EMBL Data Library, June 1998  
 A:Description: *Arabidopsis thaliana* chromosome II BAC F13M22 genomic sequence.  
 A:Reference number: Z14677  
 A:Accession: T02518  
 A:Status: translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-825 <ROU>  
 A:Cross-references: EMBL:AC004684; NID:g3236234; PID:g3236235  
 A:Experimental source: cultivar Columbia  
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tall  
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Vent  
 Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.  
 A:Reference number: A84420; MUID:20083487; PMID:10617197  
 A:Accession: G84793  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-825 <STO>  
 A:Cross-references: GB:AE002093; NID:g3236235; PIDN:AAC33623.1; GSPDB:GN00139  
 C:Genetics:  
 A:Gene: F13M22.2; At2g37520  
 A:Map position: 2  
 A:Introns: 189/3; 212/3; 282/3; 304/1; 366/2; 388/2; 415/3; 433/2; 489/1; 550/2; 56  
 C:Superfamily: *Arabidopsis thaliana* hypothetical protein F4P12.380

Query Match 52.1%; Score 38; DB 2; Length 825;  
 Best Local Similarity 50.0%; Pred. No. 74;  
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 RGTKCLRKXIPR 14  
 :|| | |||:

Db 802 KGTSMLEKKVVK 813

# RESULT 13

C86246  
hypothetical protein [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: C86246  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: C86246  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-341 <STO>  
A:Cross-references: GB:AE005172; NID:95734740; PIDN:RAD50005.1; GSPDB:GN00141  
C:Genetics:  
A:Map position: 1

Query Match 51.4%; Score 37.5; DB 2; Length 341;  
Best Local Similarity 61.5%; Pred. No. 42;  
Matches 8; Conservative 2; Mismatches 0; Indels 3; Gaps 1;

QY 3 RGT---KCLRKKI 12  
:|||:||||:  
Db 186 KGTTSKCLRKKV 198

# RESULT 14

C71554  
probable polysaccharide hydrolase - Chlamydia trachomatis (serotype D, strain UW3/Cx)  
C:Species: Chlamydia trachomatis  
C:Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 08-Oct-1999  
C:Accession: C71554  
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell,  
Science 282, 754-759, 1998  
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia tra  
A:Reference number: A71570; MUID:9900809; PMID:9784136  
A:Accession: C71554  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-283 <ARN>  
A:Cross-references: GB:AE001286; GB:AE001273; NID:93328516; PIDN:AAC67718.1; PID:9332852  
A:Experimental source: serotype D, strain UW-3/Cx  
C:Genetics:  
A:Gene: ydho

Query Match 50.7%; Score 37; DB 2; Length 283;  
Best Local Similarity 45.5%; Pred. No. 44;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 4 GTCCLRKKIPR 14  
:||:|:|:  
Db 157 GGCIIHKQLPR 167

# RESULT 15

T49514  
hypothetical protein B14D6.660 [imported] - Neurospora crassa  
C:Species: Neurospora crassa  
C:Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000  
C:Accession: T49514  
R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,  
submitted to the Protein Sequence Database, May 2000

A:Reference number: 225022  
A:Accession: T49514  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-327 <SCH>  
A:Cross-references: EMBL:AL356173; GSPDB:GN00116; NCSP:B14D6.660  
A:Experimental source: BAC clone B14D6; strain OR74A  
C:Genetics:  
A:Gene: NCSP:B14D6.660  
A:Map position: 6

Query Match 50.7%; Score 37; DB 2; Length 327;  
Best Local Similarity 54.5%; Pred. No. 50;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 RGTCLRKKIP 13  
||:||||:  
Db 276 RGPOCLRRRPP 286

Search completed: July 8, 2003, 16:29:27  
Job time : 16.1795 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2003, 16:16:30 ; Search time 6.82051 Seconds  
(without alignments)  
85.136 Million cell updates/sec

Title: US-09-824-647-5

Perfect score: 73

Sequence: 1 SARGTKCLRKIPR 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	73	100.0	589	1 GRN_MOUSE	P28798 mus musculus
2	65	89.0	588	1 GRN_HAT	P23785 r granulins
3	58	79.5	593	1 GRN_HUMAN	P28799 h granulins
4	40	54.8	555	1 SYO_YERPE	Q8zdd9 yersinia pe
5	39	53.4	67	1 BD03_HUMAN	P81534 homo sapien
6	39	53.4	424	1 YC06_METJA	Q58603 methanococc
7	38	52.1	365	1 LMCL_MOUSE	Q8veel mus musculus
8	38	52.1	461	1 TRPC_BUCSC	Q44603 buchnera ap
9	38	52.1	821	1 CTRL_ARATH	Q05609 arabidopsis
10	37	50.7	542	1 PGMO_AGRTU	P39671 agrobacteri
11	37	50.7	550	1 Y436_METJA	Q57878 methanococc
12	36	49.3	527	1 DRTS_TRYBB	Q27783 trypanosoma
13	36	49.3	1300	1 Y338_MYCPN	P75296 mycoplasma
14	36	49.3	1668	1 DPOL_TREHY	Q9hh05 thermococcu
15	35.5	48.6	591	1 GRN_CAVPO	P28797 cavia porce
16	35	47.9	41	1 BAXC_HUMAN	Q07815 homo sapien
17	35	47.9	167	1 YFCN_HAEIN	P44126 haemophilus
18	35	47.9	270	1 RL7_MOUSE	P14148 mus musculus
19	35	47.9	424	1 CYAA_STIAU	P40137 stigmatella
20	35	47.9	533	1 ALG6_ARATH	Q9ff17 arabidopsis
21	35	47.9	608	1 KU70_MOUSE	P23475 mus musculus
22	35	47.9	627	1 CACP_MOUSE	P47934 mus musculus
23	35	47.9	632	1 PTMA_BUCAL	P57635 buchnera ap
24	35	47.9	649	1 PTMA_VIBCH	Q9kkq7 vibrio chol
25	35	47.9	886	1 LEUR_YEAST	P08638 saccharomyc
26	34.5	47.3	1087	1 AKR3_RABIT	Q28628 oryctolagus
27	34	46.6	70	1 RK28_CYPAP	P48129 cyanophora
28	34	46.6	176	1 V6B_CVFE3	P33467 feline ente
29	34	46.6	201	1 RECR_TREPA	Q83969 treponema p
30	34	46.6	208	1 KGUA_MYCTU	P71659 mycobacteri
31	34	46.6	261	1 BEND_ACICA	P07772 a cis-1,2-d
32	34	46.6	338	1 REG2_YEAST	P38232 saccharomyc
33	34	46.6	354	1 NOV_MOUSE	Q64299 mus musculus

#### RESULT 1

ID	GRN_MOUSE	STANDARD;	PRT;	589 AA.
AC	P28798;			
DT	01-DEC-1992 (Rel. 24, Created)			
DT	01-OCT-1994 (Rel. 30, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Granulins precursor (Acrogranin) [Contains: Granulin 1; Granulin 2; Granulin 3; Granulin 4; Granulin 5; Granulin 6; Granulin 7].			
DE	GRN.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
EX	MEDLINE=93245991; PubMed=8482392;			
RA	Baba T., Nemoto H., Watanabe K., Arai Y., Gerton G.L.;			
RT	"Exon/Intron organization of the gene encoding the mouse			
RT	epithelin/granulin precursor (acrogranin).";			
RL	FEBS Lett. 322:89-94(1993).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Kidney;			
RA	MEDLINE=92317004; PubMed=1618805;			
RA	Plowman G.D., Green I.M., Neubauer M.G., Buckley S., McDonald V.L.,			
RT	Todaro G.I., Shoyab M.;			
RT	"The epithelin precursor encodes two proteins with opposing			
RT	activities on epithelial cell growth.";			
RL	J. Biol. Chem. 267:13073-13078(1992).			
CC	- - FUNCTION: GRANULINS HAVE POSSIBLE CYTOKINE-LIKE ACTIVITY. THEY MAY			
CC	PLAY A ROLE IN INFLAMMATION, WOUND REPAIR, AND TISSUE REMODELING.			
CC	- - TISSUE SPECIFICITY: UBIQUITOUS.			
CC	- - PTM: GRANULINS ARE DISULFIDE BRIDGED.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch)			
CC	-----			
DR	EMBL; D16195; BAA03736.1; -			
DR	EMBL; M86736; AAA37191.1; -			
DR	EMBL; X62321; CAA44197.1; -			
DR	MGD; MGI:95832; Grn.			
DR	InterPro; IPR000118; Granulin.			
DR	Pfam; PF00396; granulin; 7.			
DR	SMART; SM00277; GRAN; 7.			
DR	PROSITE; PS00799; GRANULINS; 7.			
KW	Cytokine; Repeat; Glycoprotein; Signal.			
FT	SIGNAL 1 17 POTENTIAL.			
FT	CHAIN 18 589			
FT	PEPTIDE 758 7113			
FT	PEPTIDE 7122 7178			
FT	GRANULIN 1.			
FT	GRANULIN 2.			

#### ALIGNMENTS

34	34	46.6	357	1	NOV_HUMAN	P48745 homo sapien
35	34	46.6	405	1	Y927_PYPAB	Q9uyv6 pyrococcus
36	34	46.6	447	1	SYN_STRPN	Q97pr0 streptococc
37	34	46.6	556	1	SYO_VIBCH	Q9kta6 vibrio chol
38	34	46.6	559	1	SYO_PASMO	P57847 pasteurella
39	34	46.6	620	1	YM20_YEAST	Q03162 saccharomyc
40	34	46.6	624	1	PTMA_PASMO	Q9c1y8 pasteurella
41	34	46.6	859	1	RPA2_METJA	Q58446 methanococc
42	34	46.6	940	1	CHRD_BRARE	O57472 brachydanio
43	34	46.6	956	1	MTN2_HUMAN	O00339 homo sapien
44	34	46.6	956	1	MTN2_MOUSE	O08746 mus musculu
45	34	46.6	1002	1	POL_HVIEL	P04589 human immun



FT PEPTIDE 205 260 GRANULIN 3.  
 FT PEPTIDE 280 334 GRANULIN 4.  
 FT PEPTIDE 362 7414 GRANULIN 5.  
 FT PEPTIDE 440 7493 GRANULIN 6.  
 FT PEPTIDE 7517 7568 GRANULIN 7.  
 FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 373 373 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 526 526 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 350 350 L -> R (IN REF. 2).  
 SQ SEQUENCE 589 AA; 63458 MW; 1DE8229C413CB787 CRC64;

Query Match 100.0%; Score 73; DB 1; Length 589;  
 Best Local Similarity 100.0%; Pred. No. 9e-06;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SARGTKLRRKKIPR 14  
 |||||  
 Db 562 SARGTKLRRKKIPR 575

RESULT 2  
 GRN\_RAT STANDARD; PRT; 588 AA.  
 ID GRN\_RAT  
 AC P23785;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Granulins precursor (Acrogranin) [Contains: Granulin 1 (Granulin G);  
 DE Granulin 2 (Granulin F); Granulin 3 (Granulin B) (Epithelin 2);  
 DE Granulin 4 (Granulin A) (Epithelin 1); Granulin 5 (Granulin C);  
 DE Granulin 6 (Granulin D); Granulin 7 (Granulin E)].  
 GN GRN.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=94082640; PubMed=8243292;  
 RA Bhandari V., Giald A., Bateman A.;  
 RA Todaro G.J., Shoyab M.;  
 FT "the epithelin precursor encodes two proteins with opposing  
 RT activities on epithelial cell growth.";  
 RL J. Biol. Chem. 267:13073-13078(1992).  
 RN [3]  
 RP SEQUENCE OF 204-225 AND 279-299.  
 RX MEDLINE=91045907; PubMed=2236009;  
 RA Shoyab M., McDonald V.L., Byles C., Todaro G.J., Plowman G.D.;  
 RT "Epithelins 1 and 2: isolation and characterization of two  
 RT cysteine-rich growth-modulating proteins.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:7912-7916(1990).  
 RN [4]  
 RP SEQUENCE OF 278-328.  
 RC TISSUE=Bone marrow;  
 RX MEDLINE=91097544; PubMed=22683320;  
 RA Bateman A., Belcourt D.R., Bennett H.P., Lazure C., Solomon S.;  
 RT "Granulins, a novel class of peptide from leukocytes.";  
 RL Biochem. Biophys. Res. Commun. 173:1161-1168(1990).  
 CC -1- FUNCTION: GRANULINS HAVE POSSIBLE CYTOKINE-LIKE ACTIVITY. THEY MAY  
 CC PLAY A ROLE IN INFLAMMATION, WOUND REPAIR, AND TISSUE REMODELING.  
 CC -1- TISSUE SPECIFICITY: UBIQUITOUS; MOST ABUNDANT IN THE SPLEEN AND  
 CC SEVERAL TISSUES OF ENDOCRINE SIGNIFICANCE.

CC -1- PTM: GRANULINS ARE DISULFIDE BRIDGED.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sfb.ch/announcement/  
 CC or send an email to license@sib-sfb.ch).  
 CC -----  
 DR EMBL; M97750; AAA16903.1; -;  
 DR EMBL; X62322; CAA44196.1; -;  
 DR PIR; A36199; A36199.  
 DR PIR; B36199; B36199.  
 DR PIR; E36698; E36698.  
 DR InterPro: IPR000118; Granulin.  
 DR Pfam: PF00396; granulin; 7.  
 DR SMART; SM00277; GRAN; 7.  
 DR PROSITE; PS00799; GRANULINS; 7.  
 KW Cytokine; Repeat; Glycoprotein; Signal.  
 FT SIGNAL 1 17 POTENTIAL.  
 FT CHAIN 18 588 ACROGRANIN.  
 FT PEPTIDE 58 113 GRANULIN 1.  
 FT PEPTIDE 122 178 GRANULIN 2.  
 FT PEPTIDE 204 259 GRANULIN 3.  
 FT PEPTIDE 278 334 GRANULIN 4.  
 FT PEPTIDE 361 413 GRANULIN 5.  
 FT PEPTIDE 438 492 GRANULIN 6.  
 FT PEPTIDE 512 567 GRANULIN 7.  
 FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 372 372 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 525 525 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 201 201 S -> PP (IN REF. 2).  
 FT CONFLICT 307 308 TK -> SB (IN REF. 4).  
 FT CONFLICT 324 324 Q -> T (IN REF. 4).  
 FT CONFLICT 388 388 M -> I (IN REF. 2).  
 SQ SEQUENCE 588 AA; 63369 MW; 113D434F7E099B31 CRC64;

Query Match 89.0%; Score 65; DB 1; Length 588;  
 Best Local Similarity 85.7%; Pred. No. 0.00026;  
 Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SARGTKLRRKKIPR 14  
 |||||  
 Db 561 SARGTKLRRKKIPR 574

RESULT 3  
 GRN\_HUMAN STANDARD; PRT; 593 AA.  
 ID GRN\_HUMAN  
 AC P28799; P23781; P23782; P23783; P23784; Q9BWE7;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Granulins precursor (Acrogranin) [Contains: Paragranulin; Granulin 1  
 DE Granulin 4; Granulin 2 (Granulin F); Granulin 3 (Granulin B);  
 DE Granulin 4 (Granulin A); Granulin 5 (Granulin C); Granulin 6 (Granulin  
 DE D); Granulin 7 (Granulin E)].  
 GN GRN.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE=Bone marrow;  
 RX MEDLINE=92179253; PubMed=1542665;  
 RA Bhandari V., Palfree R.G.E., Bateman A.;  
 RT "Isolation and sequence of the granulin precursor cDNA from human  
 RT bone marrow reveals tandem cysteine-rich granulin domains.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:1715-1719(1992).  
 RN [2]  
 RP REVISIONS, SEQUENCE FROM N.A.

RX MEDLINE-93038704; PubMed-1417868;  
 RA Bhandari V., Bateman A.;  
 RT "Structure and chromosomal location of the human granulin gene.";  
 RL Biochem. Biophys. Res. Commun. 188:57-63(1992).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Kidney;  
 RX MEDLINE-92317004; PubMed-1618805;  
 RA Plozman G.D., Green I.M., Neubauer M.G., Buckley S.D., McDonald V.L.,  
 RA Todaro G.J., Shoyab M.;  
 RT "The epithelin precursor encodes two proteins with opposing activities  
 on epithelial cell growth.";  
 RL J. Biol. Chem. 267:13073-13078(1992).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Brain;  
 RA Yu.W., Gibbs R.A.;  
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
 RC TISSUE=Cervix, and Lung;  
 RA Strausberg R.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE OF 206-233; 281-336; 364-396 AND 442-447.  
 RC TISSUE=Leukocyte;  
 RX MEDLINE-91097544; PubMed-2268320;  
 RA Bateman A., Belcourt D.R., Bennett H.P., Lazure C., Solomon S.;  
 RT Granulins, a novel class of peptide from leukocytes.";  
 RL Biochem. Biophys. Res. Commun. 173:1161-1168(1990).  
 CC -!- FUNCTION: GRANULINS HAVE POSSIBLE CYTOKINE-LIKE ACTIVITY. THEY MAY  
 PLAY A ROLE IN INFLAMMATION, WOUND REPAIR, AND TISSUE REMODELING.  
 CC -!- FUNCTION: GRANULIN A PROMOTES PROLIFERATION OF THE EPITHELIAL  
 CELL LINE A431 IN CULTURE WHILE GRANULIN B ACTS AS AN ANTAGONIST  
 TO GRANULIN A, INHIBITING THE GROWTH.  
 CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; may be  
 produced by alternative splicing.  
 CC -!- TISSUE SPECIFICITY: IN MYELOGENOUS LEUKEMIC CELL LINES OF  
 FIBROBLASTS, AND VERY STRONGLY IN EPITHELIAL CELL LINES. PRESENT  
 IN INFLAMMATORY CELLS AND BONE MARROW. HIGHEST LEVELS IN KIDNEY.  
 CC -!- PTM: GRANULINS ARE DISULFIDE BRIDGED.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M75161; AAA58617.1; ALT\_SEQ.  
 DR EMBL; X62320; CAA44196.1; -  
 DR EMBL; AF055008; AAC09359.1; -  
 DR EMBL; BC000324; AAH00324.1; -  
 DR EMBL; BC010577; AAH10577.1; -  
 DR PIR; A38118; GYHU  
 DR PIR; D36698; D36698.  
 DR PIR; JCI284; JCI284.  
 DR Genew; HGNC:4601; GRN.  
 DR MIM; 138945; -  
 DR InterPro; IPR000118; Granulin.  
 DR Pfam; PF00396; granulin; 7.  
 DR SMART; SM00277; GRAN; 7.  
 DR PROSITE; PS00799; GRANULINS; 7.  
 KW Cytokine; Repeat; Glycoprotein; Signal; Alternative splicing;  
 KW Polymorphism.  
 FT SIGNAL 1 17 POTENTIAL.  
 FT CHAIN 18 593 ACROGRANULIN.  
 FT PEPTIDE 18 747 PARAGRANULIN.  
 FT PEPTIDE 758 7113 GRANULIN 1.  
 FT PEPTIDE 7123 7179 GRANULIN 2.  
 FT PEPTIDE 206 261 GRANULIN 3.

FT PEPTIDE 281 336 GRANULIN 4.  
 FT PEPTIDE 364 7417 GRANULIN 5.  
 FT PEPTIDE 442 7496 GRANULIN 6.  
 FT PEPTIDE 7518 7573 GRANULIN 7.  
 FT CARBOHYD 118 118 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 236 236 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 265 265 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 368 368 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 530 530 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARSPLIC 377 531 MISSING (IN ISOFORM 2).  
 FT VARIANT 454 454 G -> Q.  
 FT /FTID-VAR\_003445.  
 FT CONFLICT 219 219 S -> H (IN REF. 6).  
 FT CONFLICT 386 386 W -> H (IN REF. 6).  
 SQ SEQUENCE 593 AA; 63473 MW; 4E402BDB16DE2819 CRC64;  
 Query Match 79.5%; Score 58; DB 1; Length 593;  
 Best Local Similarity 71.4%; Pred. No. 0.005;  
 Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 SARGTKCLRKPKPR 14  
 :|||||: :  
 DB 566 AARGTKCLRKPR 579  
 :|||||: :  
 RESULT 4  
 ID SYQ\_YERPE STANDARD; PRT; 555 AA.  
 AC Q8ZDD9;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DE 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Glutamyl-tRNA synthetase (EC 6.1.1.18) (Glutamine--tRNA ligase)  
 DE (GlnRS).  
 DE Glns OR YPO2630.  
 OS Yersinia pestis.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Yersinia.  
 OC NCBI\_TaxID=632;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=CO-92 / Biovar Orientalis;  
 RX MEDLINE-21470413; PubMed-11586360;  
 RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,  
 RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,  
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,  
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,  
 RA Feilwell T., Hamlin N., Holroyd S., Jagers K., Karlyshev A.V.,  
 RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,  
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;  
 RA "Genome sequence of Yersinia pestis, the causative agent of plague.";  
 RT Nature 413:523-527(2001).  
 RL Nature 413:523-527(2001).  
 CC -!- CATALYTIC ACTIVITY: ATP + L-glutamine + tRNA(Gln) -> AMP +  
 diphosphate + L-glutamyl-tRNA(Gln).  
 CC -!- SUBUNIT: Monomer (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AJ414153; CAC92873.1; -  
 DR InterPro; IPR004514; Glns.  
 DR InterPro; IPR000924; Glu\_tRNA-synt\_1c.  
 DR InterPro; IPR001412; tRNA-synt\_1.  
 DR Pfam; PF00749; tRNA-synt\_1c; 1.  
 DR TIGRFAMS; TIGR00440; glns; 1.  
 DR PROSITE; PS00178; AA\_TRNA\_LIGASE\_I; 1.

KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
KW Complete proteome.  
FT SITE 34 44 "HIGH" REGION.  
FT SITE 268 272 "KWSKS" REGION.  
FT BINDING 271 271 ATP (BY SIMILARITY).  
SQ SEQUENCE 555 AA; 63750 MW; AD3C68D6F56C0B5 CRC64;  
  
Query Match 54.8%; Score 40; DB 1; Length 555;  
Best Local Similarity 72.7%; Pred. No. 9.2;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 2 ARGTKLKRKI 12  
| | | | |  
DB 167 AEGTACLRKI 177  
  
RESULT 5  
BD03\_HUMAN STANDARD; PRT; 67 AA.  
AC P81534; Q9NPF6;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Beta-defensin 3 precursor (BD-3) (Defensin like protein).  
GN DEFB3 OR BD3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
SEQUENCE FROM N.A., SEQUENCE OF 23-67, FUNCTION, TISSUE SPECIFICITY,  
RP INDUCTION, AND MASS SPECTROMETRY.  
RP TISSUE-Keratinocytes, Tracheal epithelium, and Lung epithelial cells;  
RX MEDLINE=21101950; PubMed=11085990;  
RA Harder J., Bartels J., Christophers E., Schroeder J.-M.;  
RT "Isolation and characterization of human beta-defensin-3, a novel  
RT human inducible peptide antibiotic."  
RL J. Biol. Chem. 276:5707-5713(2001).  
RN [2]  
SEQUENCE FROM N.A., AND CHARACTERIZATION.  
RP PubMed=11702237;  
RA Garcia J.-R., Jaumann F., Schulz S., Krause A., Rodriguez-Jimenez J.,  
RA Forssmann U., Adermann K., Kluever E., Vogelmeier C., Becker D.,  
RA Hedrich R., Forssmann W.-G., Bals R.;  
RT "Identification of a novel, multifunctional beta-defensin (human  
RT beta-defensin 3) with specific antimicrobial activity. Its  
RT interaction with plasma membranes of *Xenopus* oocytes and the  
RT induction of macrophage chemotaxis."  
RL Cell Tissue Res. 306:257-264(2001).  
RN [3]  
SEQUENCE FROM N.A.  
RX MEDLINE=21125233; PubMed=11223260;  
RA Jia H.P., Schutte B.C., Schudy A., Linzmeier R., Guthmiller J.M.,  
RA Johnson G.K., Tack B.F., Mitros J.P., Rosenthal A., Ganz T.,  
RA McCray F.B. Jr.;  
RT "Discovery of new human defensins using a genomics-based approach."  
RL Gene 263:211-218(2001).  
RN [4]  
SEQUENCE FROM N.A.  
RA Imai Y.;  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
RN [5]  
SEQUENCE FROM N.A.  
RA Adler D.A., Diamond G., Sheppard P., Holloway J., Presnell S.,  
RA Jaspers S., Whitmore T., Fox B., Gosink J., Rixon M., Gao Z.,  
RA Haldeman B., O'Hara P.;  
RT "EST and genomic database mining yield novel human and mouse  
RT beta-defensins."  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: EXHIBITS ANTIMICROBIAL ACTIVITY AGAINST GRAM-POSITIVE  
CC BACTERIA *S. AUREUS* AND *S. PYOGENES*, GRAM-NEGATIVE BACTERIA  
CC *P. AERUGINOSA* AND *E. COLI* AND THE YEAST *C. ALBICANS*. KILLS  
CC MULTIRESISTANT *S. AUREUS* AND VANCOMYCIN-RESISTENT *E. FAECIUM*. NO

CC SIGNIFICANT HEMOLYTIC ACTIVITY WAS OBSERVED.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN SKIN AND TONSILS, AND TO A  
CC LESSER EXTENT IN TRACHEA, UTERUS, KIDNEY, THYMUS, ADENOID, PHARYNX  
CC AND TONGUE. LOW EXPRESSION IN SALIVARY GLAND, BONE MARROW, COLON,  
CC STOMACH, POLYP AND LARYNX. NO EXPRESSION IN SMALL INTESTINE.  
CC -1- INDUCTION: BY INFECTION OF BACTERIA AND BY INTERFERON GAMMA.  
CC -1- MASS SPECTROMETRY: MW=5154.59; METHOD-Electrospray; RANGE=23-67.  
CC -1- SIMILARITY: BELONGS TO THE BETA-DEFENSIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by not-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC  
CC EMBL; AJ237673; CAC03097.1; -  
CC EMBL; AF295370; AAG02237.1; -  
CC EMBL; AF217245; AAF73853.1; -  
CC EMBL; AB037972; BAB40572.1; -  
CC EMBL; AF301470; AAG22030.1; -  
CC GenBank; HGNC:15967; DEFB103.  
CC MIM; 606611; -  
CC InterPro; IPR001855; Defensin\_beta.  
CC Pfam; PF00711; Defensin\_beta; 1.  
KW Antibiotic; Signal. 22  
FT SIGNAL 1 22  
FT CHAIN 23 67 BETA-DEFENSIN 3.  
FT DISULFID 33 62 BY SIMILARITY.  
FT DISULFID 40 55 BY SIMILARITY.  
FT DISULFID 45 63 BY SIMILARITY.  
SQ SEQUENCE 67 AA; 7697 MW; 54266DE1C90D4B65 CRC64;  
  
Query Match 53.4%; Score 39; DB 1; Length 67;  
Best Local Similarity 63.6%; Pred. No. 1.9;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 SARGTKLRKK 11  
| | | | |  
DB 56 STRGRKCRKK 66  
  
RESULT 6  
YC06\_METJA STANDARD; PRT; 424 AA.  
ID YC06\_METJA  
AC Q58603;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein MJ1206.  
GN MJ1206.  
OS Methanococcus jannaschii.  
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
OC Methanocaldococcaceae; Methanocaldococcus.  
OX NCBI\_TaxID=2190;  
RN [1]  
SEQUENCE FROM N.A.  
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;  
RX MEDLINE=96337999; PubMed=8688087;  
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,  
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,  
RA Otterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
RT "Complete genome sequence of the methanogenic archaeon, *Methanococcus*  
RT *jannaschii*."  
RL Science 273:1058-1073(1996).  
CC -1- SIMILARITY: BELONGS TO THE UPF0095 FAMILY.

```
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U67562; RAB99210.1;
DR TIGR; MJ1206;
DR InterPro; IPR002936; DNAPrim_toprim.
DR Pfam; PF01751; TOPRIM; 1.
DR SMART; SM00493; TOPRIM; 1.
DR KX Hypothetical protein; Complete proteome.
DR KX SM00493; TOPRIM; 1.
SQ SEQUENCE 424 AA; 47388 MW; C55BC7A05B0EB2C3 CRC64;
Query Match 53.4%; Score 39; DB 1; Length 424;
Best Local Similarity 87.5%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 6 KCLRSKIP 13
DB 259 KCLRSKIP 266
|||||
RESULT 7
LML_MOUSE STANDARD; PRT; 365 AA.
AC Q8VEEL;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE LIM and cysteine-rich domains protein 1.
GN LMD1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 2 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC
CC IONS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; BC019124; AAH19124.1;
DR InterPro; IPR001781; LIM.
DR Pfam; PF00412; LIM; 2.
DR ProDom; PD000094; LIM; 2.
DR SMART; SM00132; LIM; 2.
DR PROSITE; PS00478; LIM_DOMAIN_1; 2.
DR PROSITE; PS00023; LIM_DOMAIN_2; 2.
DR LIM domain; Metal-binding; Zinc; Repeat.
FT DOMAIN 28 58 CYS-RICH.
FT DOMAIN 241 306 LIM 1.
FT DOMAIN 307 365 LIM 2.
SQ SEQUENCE 365 AA; 40996 MW; DB8AA836364E47D3 CRC64;
Query Match 52.1%; Score 38; DB 1; Length 365;
Best Local Similarity 72.7%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 SARGTKLRKK 11
|||||
```

```
DB 22 SARGVACLCK 32
RESULT 8
TRPC_BUCSC STANDARD; PRT; 461 AA.
AC Q44603;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tryptophan biosynthesis protein trpCF [includes: Indole-3-glycerol
DE phosphate synthase (EC 4.1.1.48) (IGPS); N-(5'-phospho-
DE ribosyl)anthranilate isomerase (EC 5.3.1.24) (PRAI)].
GN TRPC OR TRPC/F.
OS Buchnera aphidicola (subsp. Schlechtendalia chinensis).
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=118110;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=95261545; PubMed=7742976;
RA Lai C.-Y., Baumann P., Moran N.A.;
RT "Genetics of the tryptophan biosynthetic pathway of the prokaryotic
RT endosymbiont (Buchnera) of the aphid Schlechtendalia chinensis.";
RL Insect Mol. Biol. 4:47-59(1995).
CC -1- FUNCTION: BIFUNCTIONAL ENZYME THAT CATALYZES TWO SEQUENTIAL STEPS
CC OF TRYPTOPHAN BIOSYNTHETIC PATHWAY. THE FIRST REACTION IS
CC CATALYZED BY THE ISOMERASE, CODED BY THE TRPF DOMAIN; THE SECOND
CC REACTION IS CATALYZED BY THE SYNTHASE, CODED BY THE TRPC DOMAIN.
CC -1- CATALYTIC ACTIVITY: N-(5-phospho-beta-D-ribosyl)-anthranilate - 1-
CC (2-carboxyphenylamino)-1-deoxy-D-ribulose 5-phosphate.
CC -1- CATALYTIC ACTIVITY: 1-(2-carboxyphenylamino)-1-deoxy-D-ribulose 5-
CC phosphate - 1-(indol-3-yl)glycerol 3-phosphate + CO(2) + H(2)O.
CC -1- PATHWAY: Tryptophan biosynthesis; third step.
CC -1- PATHWAY: Tryptophan biosynthesis; fourth step.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE TRPC FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE TRPF FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U09185; AAA92795.1;
DR HSP; P00909; IPII.
DR InterPro; IPR001468; IGPS.
DR InterPro; IPR001240; PRAI.
DR Pfam; PF00218; IGPS; 1.
DR Pfam; PF00697; PRAI; 1.
DR ProDom; PD001511; IGPS; 1.
DR PROSITE; PS00614; IGPS; 1.
DR Tryptophan biosynthesis; Isomerase; Lyase; Multifunctional enzyme;
KW Decarboxylase.
FT DOMAIN 1 258 INDOL-3-GLYCEROL PHOSPHATE SYNTHASE.
FT DOMAIN 259 461 N-(5'-PHOSPHORIBOSYL)ANTHRANILATE
FT ISOMERASE.
SQ SEQUENCE 461 AA; 52783 MW; F202C81D6419D68D CRC64;
Query Match 52.1%; Score 38; DB 1; Length 461;
Best Local Similarity 75.0%; Pred. No. 18;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 7 CLRKIPR 14
DB 346 CLRKIPR 353
|||||
RESULT 9
CTRL_ARATH STANDARD; PRT; 821 AA.
AC Q05609;
```

DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Serine/threonine-protein kinase CTRL (EC 2.7.1.37).  
 GN CTRL OR AF5G03730 OR F17C15\_150.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia; TISSUE=Seedling;  
 RX MEDLINE=93161417; PubMed=8431946;  
 RA Kleber J.J., Rothenberg M., Roman G., Feldmann K.A., Ecker J.R.;  
 FT "CTRL, a negative regulator of the ethylene response pathway in  
 RT Arabidopsis, encodes a member of the raf family of protein kinases."  
 RL Cell 72:427-441(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=21016721; PubMed=11130714;  
 RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,  
 RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,  
 RA Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakayama S.,  
 RA Nakazaki N., Naruo K., Okumura S., Shinpo S., Takeuchi C., Wada T.,  
 RA Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,  
 RA Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,  
 RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,  
 RA Stoneking T., Pepin K., Speth J., Sekhon M., Armstrong J., Becker M.,  
 RA Belter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,  
 RA Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P.,  
 RA Leonard S., Meyer R., Mulvaney E., Ozersky P., Riley A., Strommatt C.,  
 RA Wagner-McPherson C., Wollam A., Yeakum M., Bell M., Dedhia N.,  
 RA Parnell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J.,  
 RA Carroff K., Toth K., King L., Bahret A., Miller B., Marra M.,  
 RA Martienssen R., McCombie K., Wilson R.K., Murphy G., Bancroft I.,  
 RA Volckaert G., Wambutt R., Duesterhoeft A., Stiekema W., Pohl T.,  
 RA Entlan K.-D., Terryn N., Hartley N., Bent E., Johnson S.,  
 RA Langham S.-A., McCullagh B., Robben J., Grymonprez B., Zimmermann W.,  
 RA Ramperger U., Wedler H., Balke K., Wedler E., Peters S.,  
 RA Van Staveren M., Dirkse W., Mooijman P., Klein Lankhorst R.,  
 RA Weitzenegger T., Bothe G., Rose M., Hauf J., Berner S., Hempel S.,  
 RA Feldpausch M., Lamberth S., Villarroel R., Gielen J., Ardiles W.,  
 RA Bents O., Lemcke K., Kolesov G., Mayer K.F.X., Rudd S., Schoof H.,  
 RA Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Fransz P.F.;  
 RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis  
 RL thaliana."  
 RL Nature 408:823-826(2000).  
 CC -1- FUNCTION: ACTS AS A NEGATIVE REGULATOR IN THE ETHYLENE RESPONSE  
 CC PATHWAY.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN BOTH SEEDLINGS AND ADULT PLANTS.  
 CC -1- MISCELLANEOUS: CTRL MUTANTS DISPLAY ETHYLENE-TREATED PHENOTYPES,  
 CC RESULTING IN PLANTS WITH SMALL, UNEXPANDED LEAVES AND WHOSE SEED  
 CC COTYLEDON GROWTH IS IMPAIRED.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC MIL/RAP SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L08789; AAA32779.1;  
 DR EMBL; L08790; AAA32780.1;  
 DR EMBL; AL162506; CAB82938.1;  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR004040; STY\_pkinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.

DR Pfam; PF00069; pkinase; 1.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00221; STYKC; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
 KW Transferrase; Serine/threonine-protein kinase; ATP-binding.  
 FT DOMAIN 65 69 POLY-GLY.  
 FT DOMAIN 135 141 POLY-GLY.  
 FT DOMAIN 551 809 PROTEIN KINASE.  
 FT NP\_BIND 557 585 ATP (BY SIMILARITY).  
 FT BINDING 578 585 ATP (BY SIMILARITY).  
 FT ACT\_SITE 676 676 E->K; IN CTRL-4; EXHIBITS ETHYLENE-  
 FT MUTAGEN 596 596 TREATED PHENOTYPE.  
 FT MUTAGEN 694 694 D->E; IN CTRL-1; EXHIBITS ETHYLENE-  
 FT TREATED PHENOTYPE.  
 SQ SEQUENCE 821 AA; 90306 MW; 2922D3DCDDCC15BC CRC64;  
 Query Match 52.1%; Score 38; DB 1; Length 821;  
 Best Local Similarity 57.1%; Pred. No. 31;  
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 SARGTKCLRRKKIPR 14  
 Db 761 AAVGFKRRLRLEIPR 774  
 RESULT 10  
 PGMU\_AGRTU STANDARD; PRT; 542 AA.  
 AC P39671;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Phosphoglucosyltransferase (EC 5.4.2.2) (Glucose phosphotransferase) (PGM).  
 GN PGM OR EXOC OR PSCA.  
 OS Agrobacterium tumefaciens.  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Rhizobium.  
 OX NCBI\_TaxID=358;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A348;  
 RX MEDLINE=95047522; PubMed=7959036;  
 RA Uttaro A.D., Uvalde R.A.;  
 RT "A chromosomal cluster of genes encoding ADP-glucose synthetase,  
 RT glycogen synthase and phosphoglucosyltransferase in Agrobacterium  
 RL tumefaciens."  
 RL Gene 150:117-122(1994).  
 RN [2]  
 RP CHARACTERIZATION.  
 RX MEDLINE=90170885; PubMed=2307661;  
 RA Uttaro A.D., Cangelosi G.A., Geremia R.A., Nester E.W.,  
 RA Uvalde R.A.;  
 RT "Biochemical characterization of avirulent exoC mutants of  
 RT Agrobacterium tumefaciens."  
 RL J. Bacteriol. 172:1640-1646(1990).  
 CC -1- FUNCTION: THIS ENZYME PARTICIPATES IN BOTH THE BREAKDOWN AND  
 CC SYNTHESIS OF GLUCOSE. REQUIRED FOR THE SYNTHESIS OF CAPSULAR  
 CC POLYSACCHARIDE AND NORMAL LIPOPOLYSACCHARIDE.  
 CC -1- CATALYTIC ACTIVITY: Alpha-D-glucose 1-phosphate = alpha-D-glucose  
 CC 6-phosphate.  
 CC -1- SIMILARITY: BELONGS TO THE PHOSPHOHEXOSE MUTASES FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AF033856; AAD03475.1; .  
DR HSP; P00949; 3PMG.  
DR InterPro: IPR001485; PG/PKM\_mutase.  
DR Pfam; PF00408; PGM\_PMM\_I; 1.  
DR Pfam; PF02878; PGM\_PMM\_II; 1.  
DR Pfam; PF02879; PGM\_PMM\_III; 1.  
DR Pfam; PF02880; PGM\_PMM\_IV; 1.  
DR PRINTS; PR00509; PGM\_PMM.  
DR PROSITE; PS00710; PGM\_PMM; 1.  
KW Isomerase; Phosphorylation.  
FT ACT\_SITE 112 112 FORMS THE PHOSPHOSERINE INTERMEDIATE  
FT ACT\_SITE 112 112 (BY SIMILARITY).  
SQ SEQUENCE 542 AA; 57802 MW; F40AD8ED482E905D CRC64;  
Query Match 50.7%; Score 37; DB 1; Length 542;  
Best Local Similarity 70.0%; Pred. No. 32;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
Qy 4 GTRCLRRKIP 13  
Db 16 GTSLRRKVP 25  
RESULT 11  
Y436\_METJA STANDARD; PRT; 550 AA.  
AC Q57878;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein MJ0436.  
GN MJ0436.  
OS Methanococcus jannaschii.  
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
OC Methanocaldococcaceae; Methanocaldococcus.  
OX NCBI\_TaxID=2190;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
RX MEDLINE=96337999; PubMed=9688087;  
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,  
RA Kervatage A.R., Dougherty B.A., Tomb J.F., Adams M.D., Reich C.I.,  
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,  
RA Otterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
jannaschii.";  
RL Science 273:1058-1073(1996).  
CC -1- SIMILARITY: BELONGS TO THE QUEUINE TRNA-RIBOSYLTRANSFERASE FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL; U67495; AAB98424.1; .  
DR TIGR; MJ0436.  
DR InterPro: IPR004804; Arcsd\_trna\_tgt.  
DR InterPro: IPR002478; PUA.  
DR InterPro: IPR002616; Que\_trna\_transf.  
DR InterPro: IPR004521; Unchar\_dom\_2.  
DR Pfam; PF01472; PUA; 1.  
DR Pfam; PF01702; TGT; 1.  
DR SMART; SM00339; PUA; 1.  
DR TIGRFAMs; TIGR00432; arcsn\_trna\_tgt; 1.  
DR TIGRFAMs; TIGR00449; tgt\_general; 1.  
DR TIGRFAMs; TIGR00451; unchar\_dom\_2; 1.

KW Hypothetical protein; Transferase; Glycosyltransferase;  
KW Complete proteome.  
SQ SEQUENCE 550 AA; 64278 MW; D88B5CCE1BA24CB2 CRC64;  
Query Match 50.7%; Score 37; DB 1; Length 550;  
Best Local Similarity 61.5%; Pred. No. 32;  
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
Qy 1 SARGTKCLRRKIP 13  
Db 457 SEKGAKLLWKKIP 469  
RESULT 12  
DRTS\_TRYBB STANDARD; PRT; 527 AA.  
AC Q27783;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Bifunctional dihydrofolate reductase-thymidylate synthase (DHFR-TS)  
DE [Includes: Dihydrofolate reductase (EC 1.5.1.3); Thymidylate synthase  
DE (EC 2.1.1.45)].  
OS Trypanosoma brucei brucei.  
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.  
OX NCBI\_TaxID=5702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=427;  
RX MEDLINE=96089823; PubMed=8538681;  
RA Gamarrro F., Yu P.L., Zhao J., Edman U., Greene P.J., Santi D.;  
RT "Trypanosoma brucei dihydrofolate reductase-thymidylate synthase:  
RT gene isolation and expression and characterization of the enzyme.";  
RL Mol. Biochem. Parasitol. 72:11-22(1995).  
CC -1- CATALYTIC ACTIVITY: 5,10-methylenetetrahydrofolate + dUMP -  
CC dihydrofolate + dTMP.  
CC -1- CATALYTIC ACTIVITY: 5,6,7,8-tetrahydrofolate + NADP(+) - 7,8-  
CC dihydrofolate + NADPH.  
CC -1- PATHWAY: Deoxyribonucleotide biosynthesis.  
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE  
CC DIHYDROFOLATE REDUCTASE FAMILY.  
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE THYMIDYLATE  
CC SYNTHASE FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL; U20781; AAA91362.1; .  
DR HSP; P13100; IC17.  
DR InterPro: IPR001796; DHFR.  
DR InterPro: IPR000398; Thymidylat\_synth.  
DR Pfam; PF00186; Dihfolate\_red; 1.  
DR Pfam; PF00303; thymidylat\_synth; 1.  
DR PRINTS; PR00070; DHFR.  
DR PRINTS; PR00108; THYDMSNTTHASE.  
DR PRODOM; PD001180; Thymidylat\_synth; 1.  
DR PROSITE; PS00075; DHFR; 1.  
DR PROSITE; PS00091; THYMIDYLATE\_SYNTHASE; 1.  
KW Multifunctional enzyme; Oxidoreductase; Transferase; NADP;  
KW Methyltransferase; Nucleotide biosynthesis; One-carbon metabolism.  
FT DOMAIN 1 242 DIHYDROFOLATE REDUCTASE.  
FT DOMAIN 243 527 THYMIDYLATE SYNTHASE.  
FT ACT\_SITE 409 409 BY SIMILARITY.  
SQ SEQUENCE 527 AA; 58804 MW; 39982FC553BF7601 CRC64;  
Query Match 49.3%; Score 36; DB 1; Length 527;  
Best Local Similarity 77.8%; Pred. No. 47;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 TKLRRKKIP 13  
Db 5 TRILRRKKIP 13

## RESULT 13

Y338\_MYCPN STANDARD; PRT: 1300 AA.  
AC P75296;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical lipoprotein MG338 homolog precursor (P02\_orf1300).  
GN MPN489 OR WP353.  
OS Mycoplasma pneumoniae.  
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=2104;  
[1]  
RN SEQUENCE FROM N.A.  
RP STRAIN-ATCC 29342 / M129;  
RX MEDLINE=97105885; PubMed=8948633;  
RA Himmelsreicht R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,  
Herrmann R.;  
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae".  
RL Nucleic Acids Res. 24:4420-4449(1996).  
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor (potential).  
CC -1- SIMILARITY: BELONGS TO THE MG307 / MG309 / MG338 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: A8000033; AAB96001.1; -  
CC PROSITE: PS00013; PROKAR\_LIPOPROTEIN; 1.  
CC KW Hypothetical protein; Lipoprotein; Membrane; Signal;  
CC Complete proteome.  
CC FT SIGNAL 1 26 POTENTIAL.  
CC FT CHAIN 27 1300 HYPOTHETICAL LIPOPROTEIN MG338 HOMOLOG.  
CC FT LIPID 27 27 N-ACYL DIGLYCERIDE (POTENTIAL).  
CC SQ SEQUENCE 1300 AA; 143063 MW; B8236A003A9879DB CRC64;

Query Match 49.3%; Score 36; DB 1; Length 1300;  
Best Local Similarity 53.8%; Pred. No. 1.1e+02;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

## QY 1 SARGTKLRRKKIP 13

Db 1251 SSQGTIRPKP 1263

## RESULT 14

DPOL\_THEY STANDARD; PRT: 1668 AA.  
AC Q9HH05;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE DNA polymerase (EC 2.7.7.7) [Contains: Endonuclease PI-ThyII (EC 3.1.1.-) (Thy pol-1 intein); Endonuclease PI-ThyI (EC 3.1.1.-) (Thy pol-2 intein)] (Fragment).  
GN POL.  
OS Thermococcus hydrothermalis.  
OC Archaea; Euryarchaeota; Thermococci; Thermococcaceae;  
OC Thermococcus.  
OX NCBI\_TaxID=46539;

RN

RP

RA

RT

RL

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

[1]

SEQUENCE FROM N.A.

Querellou J.J.E., Cambon M.A., Lesongeur F.O., Barbier G.;

"Thermococcales taxonomy and phylogeny based on the comparative use of

16S rDNA, 16S-23S rDNA intergenic spacer and family B DNA polymerase

genes";

Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.

[2]

CHARACTERIZATION OF PI-THYI.

MEDLINE=20512590; PubMed=11058140;

Saves I., Eleaume H., Dietrich J., Masson J.-M.;

"The Thy pol-2 intein of Thermococcus hydrothermalis is an

isoform of PI-ThyI and PI-ThyII endonucleases.";

Nucleic Acids Res. 28:4391-4396(2000).

CC -1- FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE

EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY (BY SIMILARITY).

CC -1- FUNCTION: PI-THYI AND PI-THYII ARE ENDONUCLEASES. PI-THYI CLEAVES

THE INTERSECTING SEQUENCE OF THE THY DNA POLYMERASE. IT REQUIRES A 21-

BP MINIMAL RECOGNITION SEQUENCE.

CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate

+ [DNA](N);

CC -1- P.TM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES

A POST-TRANSLATIONAL EXCISION OF THE INTERVENING REGION (INTEIN)

FOLLOWED BY PEPTIDE LIGATION (POTENTIAL).

CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.

CC -1- SIMILARITY: IN THE INTEIN SECTION; BELONGS TO THE HOMING

ENDONUCLEASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
CC EMBL: AJ245819; CAC18555.1; -  
CC HSSP: P56689; ITGO.  
CC REBASE: 4832; PI-THYI.  
CC InterPro: IPR002064; DNA\_POL\_B.  
CC InterPro: IPR003586; Hedgehog\_hintC.  
CC InterPro: IPR003587; Hedgehog\_hintN.  
CC InterPro: IPR002203; Intein.  
CC InterPro: IPR004042; Intein\_endonuc.  
CC InterPro: IPR004578; Pol2.  
CC Pfam: PF00136; DNA\_POL\_B\_1.  
CC Pfam: PF03104; DNA\_POL\_B\_exo; 1.  
CC PRINTS: PR00106; DNAPOLB.  
CC PRINTS: PR00379; Intein.  
CC SMART: SM00305; HintC; 2.  
CC SMART: SM00306; HintN; 2.  
CC SMART: SM00486; POLB; 1.  
CC TIGRFAMs: TIGR00592; pol2; 2.  
CC PROSITE: PS00116; DNA\_POLYMERASE\_B; FALSE\_NEG.  
CC PROSITE: PS00818; INTEIN\_C\_TER; 2.  
CC PROSITE: PS00819; INTEIN\_ENDONUCLEASE; 2.  
CC PROSITE: PS00817; INTEIN\_N\_TER; 2.  
CC Transferrase: DNA-directed DNA polymerase; DNA replication;  
CC DNA-binding; Hydrolase; Nuclease; Exonuclease; Multifunctional enzyme;  
CC Protein splicing; Autocatalytic cleavage; Endonuclease; Intronic homing.  
FT CHAIN 1 458 DNA POLYMERASE, 1ST PART.  
FT CHAIN 459 995 ENDONUCLEASE PI-THYI (POTENTIAL).  
FT CHAIN 996 1044 DNA POLYMERASE, 2ND PART.  
FT CHAIN 1045 1433 ENDONUCLEASE PI-THYII (POTENTIAL).  
FT CHAIN 1434 1668 DNA POLYMERASE, 3RD PART.  
SQ SEQUENCE 1668 AA; 193319 MW; 5EEB805E7A71C8 CRC64;

Query Match 49.3%; Score 36; DB 1; Length 1668;

Best Local Similarity 61.5%; Pred. No. 1.4e+02;

Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 SARGTKLRRKKIP 13

```
DB      1294  SADGTVSLRKGVP 1306

RESULT 15
GRN_CAVPO
ID GRN_CAVPO STANDARD; PRT; 591 AA..
AC P28797;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Granulins precursor (Acrogranin) [Contains: Granulin 1; Granulin 2;
DE Granulin 3; Granulin 4; Granulin 5; Granulin 6; Granulin 7]
DE (Fragment);
GN GRN.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID-10141;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 4-21.
RC TISSUE-Testis;
RX MEDLINE-93228994; PubMed-8471244;
RA Baba T., Hoff H.B. III, Nemoto H., Lee H., Orth J., Arai Y.,
RA Gerton G.L.;
RT "Acrogranin, an acrosomal cysteine-rich glycoprotein, is the
RT precursor of the growth-modulating peptides, granulins, and
RT epithelins, and is expressed in somatic as well as male germ cells.";
RL Mol. Reprod. Dev 34:233-243(1993).
CC -!- FUNCTION: GRANULINS HAVE POSSIBLE CYTOKINE-LIKE ACTIVITY. THEY MAY
CC PLAY A ROLE IN INFLAMMATION, WOUND REPAIR, AND TISSUE REMODELING.
CC -!- TISSUE SPECIFICITY: UBIQUITOUS.
CC -!- PTM: GRANULINS ARE DISULFIDE BRIDGED.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M86735; AAA37030.1; -.
DR InterPro; IPR000118; Granulin.
DR Pfam; PF00396; granulin; 7.
DR SMART; SM00277; GRAN; 6.
DR PROSITE; PS00799; GRANULINS; 6.
KW Cytokine; Repeat; Signal; Glycoprotein.
FT NON_TER 1
FT SIGNAL <1 3
FT CHAIN 4 591 ACROGRANIN.
FT PEPTIDE 253 798 GRANULIN 1.
FT PEPTIDE 7111 7164 GRANULIN 2.
FT PEPTIDE 206 261 GRANULIN 3.
FT PEPTIDE 281 335 GRANULIN 4.
FT PEPTIDE 7376 7415 GRANULIN 5.
FT PEPTIDE 439 7494 GRANULIN 6.
FT PEPTIDE 7518 7569 GRANULIN 7.
FT CARBOHYD 183 183 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 591 AA; 62586 MW; EE/C9FC8F21CB8A1 CRC64;

Query Match 48.6%; Score 35.5; DB 1; Length 591;
Best Local Similarity 35.0%; Pred. No. 65;
Matches 7; Conservative 5; Mismatches 1; Indels 7; Gaps 1;

QY 2 ARGTCRLRKK-----IPR 14
:|:|:|:|
DB 564 SQGTRCVHKHKSLLHWDLSLPR 583

Search completed: July 8, 2003, 16:25:03
Job time : 8.82051 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2003, 16:17:10 ; Search time 30.1538 Seconds  
(without alignments)  
95.665 Million cell updates/sec

Title: US-09-824-647-5

Perfect score: 73

Sequence: 1 SARGTCLRRKKIPR 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_virus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaea.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68	93.2	589	11 Q9D2V3	Q9D2V3 mus musculus
2	58	79.5	413	4 Q9H8S1	Q9H8S1 homo sapien
3	41	56.2	452	17 Q58811	Q58811 methanococc
4	40	54.8	205	12 Q8Q6G3	Q8Q6G3 ectocarpus
5	40	54.8	265	11 Q9DCP5	Q9DCP5 mus musculus
6	40	54.8	278	11 Q9QXT5	Q9QXT5 mus musculus
7	40	54.8	289	11 Q9JK23	Q9JK23 mus musculus
8	40	54.8	355	2 Q8VUF4	Q8VUF4 azoarcus ev
9	40	54.8	555	16 Q8ZBD9	Q8ZBD9 yersinia pe
10	39	53.4	124	5 Q8SV43	Q8SV43 encephalito
11	39	53.4	247	11 Q9DAG5	Q9DAG5 mus musculus
12	39	53.4	252	4 Q9Y3K4	Q9Y3K4 homo sapien
13	39	53.4	254	8 Q8WLD8	Q8WLD8 acridocarpu
14	39	53.4	288	5 Q22385	Q22385 caenorhabdi
15	39	53.4	452	11 Q912K0	Q912K0 mus musculus
16	39	53.4	777	10 Q9C789	Q9C789 arabidopsis

17	39	53.4	942	16 P74081	P74081 synchocyst
18	39	53.4	1123	5 Q9W3W0	Q9W3W0 drosophila
19	38	52.1	99	12 Q91FT0	Q91FT0 chilo iride
20	38	52.1	175	10 Q8RXS1	Q8RXS1 arabidopsis
21	38	52.1	182	2 Q9Z6G5	Q9Z6G5 zymomonas m
22	38	52.1	190	8 Q8WL97	Q8WL97 hiptage ben
23	38	52.1	207	8 Q8WL93	Q8WL93 hiraeta smil
24	38	52.1	213	8 Q8WL92	Q8WL92 excentraden
25	38	52.1	223	8 Q8WLB5	Q8WLB5 diplopterys
26	38	52.1	228	8 Q8WLD6	Q8WLD6 brachylopho
27	38	52.1	232	8 Q8WLA3	Q8WLA3 mezia arauj
28	38	52.1	238	8 Q8WLB8	Q8WLB8 lasiocarpus
29	38	52.1	244	8 Q8WLB4	Q8WLB4 ryssocpterys
30	38	52.1	245	8 Q8WLC7	Q8WLC7 peregrina l
31	38	52.1	245	8 Q8WLB0	Q8WLB0 heladena mu
32	38	52.1	245	8 Q8WL76	Q8WL76 lophopterys
33	38	52.1	246	8 Q8WLB4	Q8WLB4 diacidia fe
34	38	52.1	246	8 Q8WLD1	Q8WLD1 malpighia e
35	38	52.1	246	8 Q8WL72	Q8WL72 aspicaarpa b
36	38	52.1	247	8 Q8WLC4	Q8WLC4 janusia lin
37	38	52.1	247	8 Q8WLB8	Q8WLB8 peixotoa gl
38	38	52.1	247	8 Q8WLB3	Q8WLB3 sphebamocca
39	38	52.1	247	8 Q8WLA5	Q8WLA5 heteroptery
40	38	52.1	247	8 Q8WL75	Q8WL75 jubelina ro
41	38	52.1	248	8 Q8WL90	Q8WL90 ectopoptery
42	38	52.1	248	8 Q8WL87	Q8WL87 ptilochoaeta
43	38	52.1	248	8 Q8WL86	Q8WL86 ptilochoaeta
44	38	52.1	248	8 Q8WL86	Q8WL86 ptilochoaeta
45	38	52.1	249	8 Q8WLD3	Q8WLD3 triaspis hy

#### ALIGNMENTS

#### RESULT 1

Q9D2V3 Q9D2V3 PRELIMINARY; PRT: 589 AA.  
AC Q9D2V3;  
DT 01-JUN-2001 (TREMREL. 17, Created)  
DT 01-JUN-2001 (TREMREL. 17, Last sequence update)  
DT 01-JUN-2002 (TREMREL. 21, Last annotation update)  
DE Adult male kidney cDNA, RIKEN full-length enriched library,  
DE clone:0610012H06, full insert sequence.  
GN GRN.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=KIDNEY;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K. I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner T., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gliss C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staibil F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
RA Hayashizaki Y.;  
RT \*Functional annotation of a full-length mouse cDNA collection.\*;  
RL Nature 409:685-690(2001).  
EMBL; AK018744; BAB31384.1; .

```
DR MGD; MGI:95832; Grn.
DR InterPro: IPR000118; Granulin.
DR InterPro: IPR001211; PhospholipaseA2.
DR Pfam: PF00396; granulin; 7.
DR SMART: SM00277; GRAN; 7.
DR PROSITE: PS00799; GRANULINS; 7.
DR PROSITE: PS00118; PA2_HIS; UNKNOWN_1.
SQ SEQUENCE 589 AA; 63405 MW; 1DE8229C413CA292 CRC64;

Query Match 93.2%; Score 68; DB 11; Length 589;
Best Local Similarity 100.0%; Pred. No. 6.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SARGTKCLRKIP 13
   |||||
DB 562 SARGTKCLRKIP 574

RESULT 2
Q9H8S1 PRELIMINARY; PRT; 413 AA.
AC Q9H8S1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE OVARC1001154 protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC TISSUE-OVARIAN CARCINOMA;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sato H., Sugano S., Shiratori A., Sudo H.,
RA Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takanashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuho Y.,
RA Ninomiya K., Iwayanagi T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK023348; BAB14535.1; -
DR InterPro: IPR000118; Granulin.
DR Pfam: PF00396; granulin; 4.
DR SMART: SM00277; GRAN; 4.
DR PROSITE: PS00799; GRANULINS; UNKNOWN_4.
DR PROSITE: PS00118; PA2_HIS; UNKNOWN_1.
SQ SEQUENCE 413 AA; 44132 MW; 0E3767A44BE314EC CRC64;

Query Match 79.5%; Score 58; DB 4; Length 413;
Best Local Similarity 71.4%; Pred. No. 0.0042;
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SARGTKCLRKIP 14
   |||||
DB 386 AARGTKCLREAPR 399

RESULT 3
Q58811 PRELIMINARY; PRT; 452 AA.
AC Q58811;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein MJ1416.
GN MJ1416.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanococcaceae; Methanococcus.
OX NCBI_TaxID=2190;

[1]
RN SEQUENCE FROM N.A.
RP STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE-96337999; PubMed-8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Goddek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Keiley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii.";
RL Science 273:1058-1073(1996).
CC -1- SIMILARITY: STRONG, TO M.THERMOAUTOTROPHICUM MTH1525, A.FULGIDUS
CC AF2363 AND P.SOLFATARICUS ORF C04046.
DR EMBL: U67582; AAB99426.1; -
DR TIGR: MJ1416; -
DR InterPro: IPR000728; AIRS_related.
DR Pfam: PF00586; AIRS; 1.
DR Pfam: PF02769; AIRS_C.1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 452 AA; 50075 MW; 8B4C364CEDB8BF92 CRC64;

Query Match 56.2%; Score 41; DB 17; Length 452;
Best Local Similarity 87.5%; Pred. No. 10;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 KCLRKKIP 13
   |||||
DB 10 RCLRKKIP 17

RESULT 4
Q8QNG3 PRELIMINARY; PRT; 205 AA.
AC Q8QNG3;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE ESV-1-227.
OS Ectocarpus siliculosus virus.
OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Phaeovirus.
OX NCBI_TaxID=37665;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=ESV-1;
RA Delaroque N., Bothe G., Pohl T., Knippers R., Mueller D.G., Boland W.;
RT "The complete nucleotide sequence of the Ectocarpus siliculosus virus
genome.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF204951; AAK14640.1; -
SQ SEQUENCE 205 AA; 24317 MW; C3A21D3650D46F32 CRC64;

Query Match 54.8%; Score 40; DB 12; Length 205;
Best Local Similarity 66.7%; Pred. No. 7.9;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SARGTKCLRKIP 12
   ||: |||||
DB 80 SAQKNKCLREKI 91

RESULT 5
Q9DCP5 PRELIMINARY; PRT; 265 AA.
AC Q9DCP5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Vascular endothelial zinc finger 1.
GN VEZF1.
```

OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=KIDNEY;  
 RX MEDLINE=21083660; PubMed=11217851;  
 RA Kawai J., Shingawa A., Shibata Y., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schrim L.M., Stauber F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
 RA Hayashizaki Y.;  
 RT \*Functional annotation of a full-length mouse cDNA collection.\*;  
 RL Nature 409:685-690(2001).  
 DR EMBL; AK002601; BAB22222.1; -;  
 DR HSSP; P00740; 1EDM.  
 DR MGD; MGI:1313291; Vezfl.  
 DR InterPro; IPR000561; EGF-like.  
 DR Pfam; PF00008; EGF\_2.  
 DR SMART; SM00181; EGF\_2.  
 DR PROSITE; PS00001; EGF-like; 2.  
 DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
 DR PROSITE; PS01186; EGF\_2; UNKNOWN\_1.  
 DR PROSITE; PS01187; EGF\_CA; 1.  
 SQ SEQUENCE 265 AA; 28665 MW; D81EC3DA884FF97E CRC64;  
 Query Match 54.8%; Score 40; DB 11; Length 265;  
 Best Local Similarity 61.5%; Pred. No. 10;  
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 OY 1 SARGTKCLRKIP 13  
 II III I I  
 DB 174 SADGTRCLSKKP 186  
 RESULT 6  
 O9QXT5  
 ID O9QXT5 PRELIMINARY; PRT; 278 AA.  
 AC O9QXT5  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE NOTCH4-like protein (Vascular endothelial zinc finger 1).  
 GN VEZF1 OR ZNED1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Sheppard P., Jelinek L., Whitmore T., Blumberg H., Lehner J.,  
 RA O'Hara P.;  
 RT \*Mus musculus homolog of HLA class III region containing NOTCH4  
 gene.\*;  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Presnell S., Gilbert T., Whitmore T., Foster D., Hart C., Lehner J.,

RA Martinez T., Hoffman R., O'Hara P.;  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LIVER;  
 RA Strausberg R.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF184973; AAF01322.1; -;  
 DR EMBL; BC024610; AAH24610.1; -;  
 DR HSSP; P00740; 1EDM.  
 DR MGD; MGI:1313291; Vezfl.  
 DR InterPro; IPR000561; EGF-like.  
 DR InterPro; IPR001881; EGF\_CA.  
 DR Pfam; PF00008; EGF\_2.  
 DR SMART; SM00181; EGF\_1.  
 DR SMART; SM00001; EGF-like; 1.  
 DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
 DR PROSITE; PS01186; EGF\_2; UNKNOWN\_1.  
 DR PROSITE; PS01187; EGF\_CA; 1.  
 DR Calcium-binding; EGF-like domain; Glycoprotein; Repeat.  
 KW SEQUENCE 278 AA; 30125 MW; 863159F567102FA7 CRC64;  
 Query Match 54.8%; Score 40; DB 11; Length 278;  
 Best Local Similarity 61.5%; Pred. No. 10;  
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 OY 1 SARGTKCLRKIP 13  
 II III I I  
 DB 174 SADGTRCLSKKP 186  
 RESULT 7  
 O9JK23  
 ID O9JK23 PRELIMINARY; PRT; 289 AA.  
 AC O9JK23  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Leucine rich protein.  
 GN DSCR2 OR MC21-LRP.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20329221; PubMed=10872820;  
 RA Vidal-Taboada J.M., Lu A., Pique M., Pons G., Gil J., Oliva R.;  
 RT \*Down syndrome critical region gene 2: expression during mouse  
 development and in human cell lines indicates a function related to  
 cell proliferation.\*;  
 RL Biochem. Biophys. Res. Commun. 272:156-163(2000).  
 DR EMBL; AJ238270; CAB87195.1; -;  
 DR MGD; MGI:1860263; DSCR2.  
 SQ SEQUENCE 289 AA; 33104 MW; FEB4F7FDDC020F06 CRC64;  
 Query Match 54.8%; Score 40; DB 11; Length 289;  
 Best Local Similarity 61.5%; Pred. No. 11;  
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 OY 1 SARGTKCLRKIP 13  
 I I I I I I I  
 DB 256 SSRSLKLVKNIP 268  
 RESULT 8  
 O8VUF4  
 ID O8VUF4 PRELIMINARY; PRT; 355 AA.  
 AC O8VUF4  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE 6-hydroxycyclohex-1-ene-1-carboxyl-CoA dehydrogenase.

GN BZDX.  
 OS Azoriscus evansii.  
 OC Bacteria; Proteobacteria; beta subdivision; Rhodocyclus group;  
 OC Azoriscus.  
 OX NCBI\_TaxID=59406;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-KB740;  
 RA Mohr K., Hermann H., Burchardt G.;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ428529; CAD21637.1; -.  
 DR InterPro; IPR002085; Adh\_zn\_family.  
 DR InterPro; IPR000345; CytC\_heme\_bind.  
 DR Pfam; PF00107; adh\_zinc; 1  
 DR PROSITE; PS00190; CYTOCHROME\_C; UNKNOWN.1.  
 SQ SEQUENCE 355 AA; 38211 MW; 5E1D930131F800A6 CRC64;  
 Query Match 54.8%; Score 40; DB 2; Length 355;  
 Best Local Similarity 46.2%; Pred. No. 13;  
 Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 SARGYKCLRKKIP 13  
 : || :|| :| :|  
 Db 99 TCGNRCILSRKP 111  
 RESULT 9  
 ID Q8ZDD9 PRELIMINARY; PRT; 555 AA.  
 AC Q8ZDD9;  
 DT 01-MAR-2002 (TREMblrel. 20, Created)  
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE Glutaminyl-tRNA synthetase (EC 6.1.1.18).  
 GN GLN OR YPO2630.  
 OS Yersinia pestis.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Yersinia.  
 OX NCBI\_TaxID=632;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CO-92 / BIOVAR ORIENTALIS;  
 RX MEDLINE=21470413; PubMed=11586360;  
 RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,  
 RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,  
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,  
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,  
 RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,  
 RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,  
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;  
 RT "Genome sequence of Yersinia pestis, the causative agent of plague."  
 RL Nature 413:523-527(2001).  
 DR EMBL; AJ414153; CAC92873.1; -.  
 DR InterPro; IPR004514; GLN.  
 DR InterPro; IPR000924; Glu\_trna-synt\_1c.  
 DR InterPro; IPR001412; trna-synt\_1.  
 DR Pfam; PF00749; trna-synt\_1c; 1.  
 DR TIGRFAMs; TIGR00440; glnS; 1.  
 DR PROSITE; PS00178; AA\_TRNA\_LIGASE\_I; 1.  
 KW Aminoacyl-tRNA synthetase; Ligase; Complete proteome.  
 SQ SEQUENCE 555 AA; 63750 MW; AD3C68D6FB56C0B5 CRC64;

Query Match 54.8%; Score 40; DB 16; Length 555;  
 Best Local Similarity 72.7%; Pred. No. 20;  
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 ARGTRKCLRKKI 12  
 : || ||| ||  
 Db 167 ABSTACLRKKI 177  
 RESULT 10  
 Q8SV43

ID Q8SV43 PRELIMINARY; PRT; 124 AA.  
 AC Q8SV43;  
 DT 01-JUN-2002 (TREMblrel. 21, Created)  
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE Hypothetical protein ECU07\_0240.  
 GN ECU07\_0240.  
 OS Encephalitozoon cuniculi.  
 OC Eukaryota; Microsporidia; Unikaryonidae; Encephalitozoon.  
 OX NCBI\_TaxID=6035;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CB-M1;  
 RA Genoscope;  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CB-M1;  
 RX MEDLINE=21576510; PubMed=11719806;  
 RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarar F.,  
 RA Prensier G., Barbe V., Peyretallade E., Brottier P., Wincker P.,  
 RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,  
 RA Weissenbach J., Vivares C.P.;  
 RT "Genome sequence and gene compaction of the eukaryote parasite  
 Encephalitozoon cuniculi.";  
 RL Nature 414:450-453(2001).  
 DR EMBL; AL590447; CAD25556.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 124 AA; 13235 MW; CBE8854640A8D841 CRC64;  
 Query Match 53.4%; Score 39; DB 5; Length 124;  
 Best Local Similarity 70.0%; Pred. No. 7.9;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 5 TKCLRKKIPR 14  
 : ||||| :  
 Db 3 TKCLRKKIEK 12  
 RESULT 11  
 ID Q9DAG5 PRELIMINARY; PRT; 247 AA.  
 AC Q9DAG5;  
 DT 01-JUN-2001 (TREMblrel. 17, Created)  
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 DE 1700011E24Rik protein.  
 GN 1700011E24Rik.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-C57BL/6J; TISSUE=TESTIS;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Stauble F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Suzuki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,

```

RA Hayashizaki Y.;
RT *Functional annotation of a full-length mouse cDNA collection.*;
RL Nature 409:685-690(2001).
DR EMBL; AK005861; BAB24281.1; -.
DR MGD; MGI:1922717; 1700011E24Rik.
SQ SEQUENCE 247 AA; 27631 MW; 5A9BBD78EE7BD9FA CRC64;

Query Match 53.4%; Score 39; DB 11; Length 247;
Best Local Similarity 60.0%; Pred. No. 15;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 TKCLRKKIPR 14
  I I I I I I
Db 194 TSCFRSKVPR 203

RESULT 12
QY3K4 PRELIMINARY; PRT; 252 AA.
ID QY3K4;
AC QY3K4;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE DJ336H9.2 (Novel AP-2 like protein) (Fragment).
GN DJ336H9.2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Peck A.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL031224; CAB41306.1; -.
DR InterPro; IPR004979; TF_AP2.
DR Pfam; PF03299; TF_AP-2; 1.
FT NON_TER 1
SQ SEQUENCE 252 AA; 27527 MW; 461BB71DC17C5980 CRC64;

Query Match 53.4%; Score 39; DB 4; Length 252;
Best Local Similarity 50.0%; Pred. No. 15;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 SARGTKCLRKKIPR 14
  I I I I I I
Db 58 SKNGGCLREKLDLR 71

RESULT 13
Q8WLD8 PRELIMINARY; PRT; 254 AA.
ID Q8WLD8;
AC Q8WLD8;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE NADH dehydrogenase subunit F (Fragment).
GN NDHF.
OS Acridocarpus natalitius.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Malpighiales; Malpighiaceae; Acridocarpaceae.
OX NCBI_TaxID=151788;
RN [1]
RP SEQUENCE FROM N.A.
RA Davis C.C., Anderson W.R., Donoghue M.J.;
RT *Phylogeny of Malpighiaceae: evidence from chloroplast ndhF and trnL-F
RT nucleotide sequences.*;
RL Am. J. Bot. 88:1830-1846(2001).
DR EMBL; AF351016; RAL35030.1; -.
DR InterPro; IPR002128; Oxidored_q1_C.
DR Pfam; PF01010; oxidored_q1_C; 1.
KW Chloroplast.

```

```

FT NON_TER 1
SQ SEQUENCE 254 AA; 29522 MW; EE5D5B653867D926 CRC64;

Query Match 53.4%; Score 39; DB 8; Length 254;
Best Local Similarity 63.6%; Pred. No. 15;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 RGTCLRKKIP 13
  I I I I I I
Db 48 QGSKLKKIP 58

RESULT 14
Q22385 PRELIMINARY; PRT; 288 AA.
ID Q22385;
AC Q22385;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE T11A5.5 protein.
GN T11A5.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA McMurray A.A.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z72515; CAA96685.1; -.
DR InterPro; IPR000896; Hemocyanin.
DR PROSITE; PS00210; HEMOCYANIN_2; UNKNOWN_1.
SQ SEQUENCE 288 AA; 32287 MW; EB1EA0EF945F9BD CRC64;

Query Match 53.4%; Score 39; DB 5; Length 288;
Best Local Similarity 70.0%; Pred. No. 17;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 GYKCLRKKIP 13
  I I I I I I
Db 145 GSKCLKKAP 154

RESULT 15
Q912K0 PRELIMINARY; PRT; 452 AA.
ID Q912K0;
AC Q912K0;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Transcription factor AP-2delta.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Zhao F., Satoda M., Licht J.D., Hayashizaki Y., Gelb B.D.;
RT "Cloning and characterization of a novel mouse AP-2 transcription
RT factor, AP-2delta, with unique DNA binding and transactivation
RT properties.";
RL J. Biol. Chem. 275:10000-10004(2000).
DR EMBL; AF421891; RAL16940.1; -.
DR InterPro; IPR004979; TF_AP2.

```

DR Pfam; PF03299; TF\_AP-2; 1.  
SQ SEQUENCE 452 AA; 49347 MW; 6722342B15CDF9EB CRC64;  
Query Match 53.4%; Score 39; DB 11; Length 452;  
Best Local Similarity 50.0%; Pred. No. 26;  
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
Qy 1 SARGTKCLRKKIPR 14  
I I :|||:|  
Db 258 SKNGGRCLREKIDR 271

Search completed: July 8, 2003, 16:28:00  
Job time : 32.1538 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2003, 16:08:55 ; Search time 47.5 seconds

(without alignments)  
53.300 Million cell updates/sec

Title: US-09-824-647-6.

Perfect score: 98

Sequence: 1 ERAPAHLSLPDQALKRDV 19

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_101002.\*  
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	98	100.0	19	AAW85480	Human GP88 autocri
2	98	100.0	413	AAW85480	Human protein sequ
3	98	100.0	593	AAW85480	Human epithelin pr
4	98	100.0	593	AAW85480	Granulin sequence.
5	98	100.0	593	AAW85480	Human GP88 autocri
6	98	100.0	593	AAW85480	Human granulin/epi
7	98	100.0	621	AAW85480	Human cancer assoc
8	63	64.3	589	AAW85480	Rat epithelin prec
9	53	54.1	179	AAW85480	Bovine epithelin p
10	50	51.0	589	AAW85480	Mouse GP88 autocri

11	50	51.0	589	23	AAE20520	Mouse granulin/epi
12	49	50.0	730	22	ABG28448	Novel human diagno
13	47	48.0	324	22	AAW84335	Human protein sequ
14	47	48.0	828	22	AAW84335	Human polypeptide
15	47	48.0	828	22	AAW84335	Human protein sequ
16	47	48.0	828	22	AAW84335	Human protein sequ
17	47	48.0	828	22	AAW84335	Human polypeptide
18	47	48.0	966	22	AAW84335	Human protein sequ
19	45	45.9	55	22	AAW84335	Propionibacterium
20	45	45.9	55	22	AAW84335	Propionibacterium
21	45	45.9	55	22	AAW84335	Propionibacterium
22	45	45.9	55	22	AAW84335	Propionibacterium
23	45	45.9	55	22	AAW84335	Propionibacterium
24	45	45.9	55	22	AAW84335	Propionibacterium
25	45	45.9	55	22	AAW84335	Propionibacterium
26	45	45.9	55	22	AAW84335	Propionibacterium
27	45	45.9	55	22	AAW84335	Propionibacterium
28	45	45.9	55	22	AAW84335	Propionibacterium
29	44	44.9	367	23	AAW84335	Propionibacterium
30	44	44.9	367	23	AAW84335	Propionibacterium
31	44	44.9	367	23	AAW84335	Propionibacterium
32	43.5	44.4	619	22	AAW84335	Propionibacterium
33	43	43.9	57	22	AAW84335	Propionibacterium
34	43	43.9	79	22	AAW84335	Propionibacterium
35	43	43.9	114	21	AAW84335	Propionibacterium
36	43	43.9	117	22	AAW84335	Propionibacterium
37	43	43.9	245	22	AAW84335	Propionibacterium
38	43	43.9	479	21	AAW84335	Propionibacterium
39	43	43.9	479	23	AAW84335	Propionibacterium
40	43	43.9	487	22	AAW84335	Propionibacterium
41	43	43.9	522	23	AAW84335	Propionibacterium
42	43	43.9	812	22	AAW84335	Propionibacterium
43	43	43.9	820	22	AAW84335	Propionibacterium
44	42.5	43.4	110	23	AAW84335	Propionibacterium
45	42	42.9	68	22	AAW84335	Propionibacterium

#### ALIGNMENTS

RESULT 1

AAW85480

ID AAW85480 standard; Peptide; 19 AA.

XX AAW85480;

XX AAW85480;

XX AAW85480;

DT 15-MAR-1999 (first entry)

XX Human GP88 autocrine growth factor antigenic peptide E19V.

DE GP88; granulin; epithelin; human; growth factor; autocrine; tumour;

XX GP88; granulin; epithelin; human; growth factor; autocrine; tumour;

KW cancer; viral infection; antagonist; therapy; diagnosis; antigen;

KW antibody.

XX Homo sapiens.

OS Homo sapiens.

XX WO9852607-A1.

XX 26-NOV-1998.

XX 22-MAY-1998; 98WO-US10555.

XX 16-DEC-1997; 97US-0991862.

XX 23-MAY-1997; 97US-0863079.

XX (SERR/) SERRERO G.

XX Serrero G;

XX WPI; 1999-045276/04.

XX Composition containing antagonist of growth factor GP88 - useful for

PT treating cancer and viral diseases and also for diagnosing disease

PT from altered GP88 expression

XX Example 8; Page 45; 86pp; English.

XX This is the amino acid sequence of peptide E19V, comprising amino  
 CC acid residues E340-V364 of human GP88 (see AAW85475). GP88 is an 88  
 CC kDa glycoprotein autocrine growth factor that is expressed in a  
 CC tightly regulated manner in normal cells, is overexpressed and  
 CC unregulated in highly tumorigenic cells derived from normal cells,  
 CC and which acts as a stringently required growth stimulator for the  
 CC tumorigenic cells. E19V was used to raise neutralising antibodies  
 CC to GP88. Antagonists to GP88, such as anti-GP88 antibodies, are  
 CC used to treat diseases associated with increased expression of  
 CC GP88, particularly cancer but also viral infections. Anti-GP88  
 CC antibodies can also be used as diagnostic reagents and to deliver  
 CC toxins or other compounds to GP88-expressing cells.

XX SQ Sequence 19 AA;

Query Match 100.0%; Score 98; DB 20; Length 19;

Best Local Similarity 100.0%; Pred. No. 1.3e-07; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKAPAHLSLPDPQALKRDV 19  
 DB 1 EKAPAHLSLPDPQALKRDV 19

RESULT 2

AA94550  
 ID AAB94550 standard; Protein; 413 AA.

AC AAB94550;

DT 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:15310.

DE Human; primer; detection; diagnosis; antisense therapy; gene therapy.

OS Homo sapiens.

PN EP1074617-A2.

XX 07-FEB-2001.

PF 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

PA Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

DR Primer sets for synthesizing polynucleotides, particularly the 5602

PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -

XX Claim 8; SEQ ID 15310; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesising polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human CDNA sequences; AAB92446 to  
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.

XX SQ Sequence 413 AA;

Query Match 100.0%; Score 98; DB 22; Length 413;

Best Local Similarity 100.0%; Pred. No. 3.7e-06; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKAPAHLSLPDPQALKRDV 19  
 DB 166 EKAPAHLSLPDPQALKRDV 184

RESULT 3

AAH14326

ID AAR14326 standard; Protein; 593 AA.

AC AAR14326;

DT 17-JAN-1992 (first entry)

DE Human epithelin precursor.

DE ET; growth regulation; inhibition; stimulation.

OS Homo sapiens.

FH Key

FT Protein

FT Location/Qualifiers

FT 1..593

FT /label= precursor.

FT /note= "claim 1, page 53"

FT 282..337

FT /label= EP-1

FT /note= "claim 2, page 53"

FT 206..262

FT /label= EP-2

FT /note= "claim 3, page 53"

FT 59..114

FT /label= EP

FT /note= "claim 4, page 53"

FT 124..180

FT /label= EP

FT /note= "claim 5, page 53"

FT 364..418

FT /label= EP

FT /note= "claim 6, page 53"

FT 442..497

FT /label= EP

FT /note= "claim 7, page 53"

FT 519..574

FT /label= EP

FT /note= "claim 8, page 55"

XX WO9115510-A.

PN 17-OCT-1991.





CC antibodies (used as antagonists, as diagnostic reagents and for  
 CC delivering toxins or other compounds to GP88-expressing cells) and  
 CC to screen for antibodies. Methods are provided for diagnosing  
 CC disease, or determining susceptibility to disease, resulting from  
 CC altered GP88 activity.

XX  
 SQ Sequence 593 AA;

Query Match 100.0%; Score 98; DB 20; Length 593;  
 Best Local Similarity 100.0%; Pred. No. 5.5e-06;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EKAPAHLSLPDPQALKRDV 19  
 |||||  
 Db 346 EKAPAHLSLPDPQALKRDV 364

RESULT 6  
 AAE20521  
 ID AAE20521 standard; Protein; 593 AA.  
 XX  
 AC AAE20521;  
 DT 01-JUL-2002 (first entry)  
 DE Human granulin/epithelin precursor (GP88) protein.  
 DE Human; granulin precursor; GP88; cytostatic; tumorigenicity; tamoxifen;  
 KW antineoplastic; antioestrogen therapy; skin cancer.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 273 /note= "Encoded by AAG"  
 FT Region 346..364  
 FT /note= "Region used as antigen to develop anti-human  
 GP88 neutralising antibody"

XX US2002025543-A1.  
 XX 28-FEB-2002.  
 XX 15-JUN-2001; 2001US-0880842.  
 XX 23-MAY-1997; 97US-0863079.  
 XX 08-DEC-1999; 99US-0456886.  
 XX (SERR/) SERRERO G.  
 XX Serrero G;  
 XX WPI; 2002-267529/31.  
 XX N-PSDB; AAD32850.  
 XX Diagnosing tumorigenicity in a human, comprising obtaining a cell  
 sample, detecting GP88 in the cells, and determining the number of GP88  
 positive cells in the sample.

XX Disclosure; Fig 9B; 50pp; English.

CC The invention relates to a method for diagnosing tumorigenicity in a  
 CC human. The method comprises obtaining a biological sample containing  
 CC cells from the patient, detecting GP88 in the cells of the sample, and  
 CC and determining the number of GP88 positive cells in the sample, and  
 CC determining the ratio of GP88 positive cells to the total number of cells  
 CC in the sample. The invention also relates to a method for determining  
 CC if a human patient is resistant to the antineoplastic effects of  
 CC antioestrogen therapy. The method is useful for diagnosing  
 CC tumorigenicity in a sample, such as blood, serum, plasma, urine, nipple  
 CC aspirate, cerebrospinal fluid, liver, kidney, breast, bone, brain, lung,  
 CC colon, or skin cancer. The method can be used to treat or prevent  
 CC re-occurrence of cancer in a patient, by administering tamoxifen if the

CC sample contains less than 10 % GP88, or less than 5 % GP88 positive  
 CC cells. The present sequence is human granulin/epithelin precursor (GP88)  
 CC protein.

XX  
 SQ Sequence 593 AA;

Query Match 100.0%; Score 98; DB 23; Length 593;  
 Best Local Similarity 100.0%; Pred. No. 5.5e-06;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EKAPAHLSLPDPQALKRDV 19  
 |||||  
 Db 346 EKAPAHLSLPDPQALKRDV 364

RESULT 7  
 AAB43971  
 ID AAB43971 standard; Protein; 621 AA.  
 XX  
 AC AAB43971;  
 DT 08-FEB-2001 (first entry)  
 DE Human cancer associated protein sequence SEQ ID NO:1416.  
 DE Human; cancer associated gene; cancer antigen; detection; cancer;  
 KW diagnosis; cytostatic; proliferative; vulnery; immunomodulator;  
 KW antiadipetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;  
 KW antidiabetic; antihypertensive; antiallergic; antibacterial; cardiant;  
 KW dermatological; neuroprotective; thrombolytic; coagulant; neotropic;  
 KW vasotropic; antipsoriatic; angiogenic; gene therapy; inflammation;  
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;  
 KW allergic reaction; graft versus host disease; organ rejection;  
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;  
 KW neurological disease; drug screening.

XX Homo sapiens.  
 OS WO2000055350-A1.  
 XX 21-SEP-2000.  
 XX 08-MAR-2000; 2000WO-US05882.  
 XX 12-MAR-1999; 99US-0124270.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX Rosen CA, Ruben SM;  
 XX WPI; 2000-587533/55.  
 XX N-PSDB; AAC78180.  
 XX Novel isolated nucleic acids comprising sequences encoding peptides  
 useful for treating or diagnosing e.g. cancer -  
 XX Claim 11; Page 2094-2096; 2352pp; English.  
 XX AAC77607 to AAC78448 encode the human cancer associated proteins given  
 CC in AAB43398 to AAB44239. The proteins can have activities based on the  
 CC tissues and cells the genes are expressed in. Example of activities  
 CC include: cytostatic; proliferative; vulnery; immunomodulator;  
 CC antiadipetic; antiasthmatic; antirheumatic; antiarthritic;  
 CC antidiabetic; antihypertensive; antiallergic; antibacterial; antiviral;  
 CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;  
 CC neotropic; vasotropic; antipsoriatic and angiogenic. The  
 CC polynucleotides and polypeptides can be used for preventing, treating or  
 CC ameliorating medical conditions and diagnosing pathological conditions.  
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from  
 CC the present invention may be used to treat immune disorders by activating  
 CC or inhibiting the proliferation, differentiation or mobilisation of  
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune  
 CC disorders, allergic reactions, graft versus host disease and organ



## CC inhibition of the

CC human. The method comprises obtaining a biological sample containing

CC cells from the patient, detecting GP88 in the cells of the sample,  
CC and determining the number of GP88 positive cells in the sample, and  
CC determining the ratio of GP88 positive cells to the total number of cells  
CC in the sample. The invention also relates to a method for determining  
CC if a human patient is resistant to the antineoplastic effects of  
CC antioestrogen therapy. The method is useful for diagnosing  
CC tumorigenicity in a sample, such as blood, serum, plasma, urine, nipple  
CC aspirate, cerebrospinal fluid, liver, kidney, breast, bone, brain, lung,  
CC colon, or skin cancer. The method can be used to treat or prevent  
CC re-occurrence of cancer in a patient, by administering tamoxifen if the  
CC sample contains less than 10 % GP88, or less than 5 % GP88 positive  
CC cells. The present sequence is mouse granulatin/epithelin precursor (GP88)  
CC protein.

XX  
SQ Sequence 589 AA;  
Query Match 51.08; Score 50; DB 23; Length 589;  
Best Local Similarity 61.18; Pred. No. 50;  
Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;  
QY 1 EKAPAHLSLPDQALKRD 18  
: | | | | | | | | | |  
Db 344 KVIAPLRPLDQILKSD 361

RESULT 12  
ABG28448  
ID ABG28448 standard; Protein: 730 AA.  
XX  
AC ABG28448;  
XX  
DT 18-FEB-2002 (first entry)  
XX  
DE Novel human diagnostic protein #28439.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.  
XX WO200175067-A2.  
XX PD 11-OCT-2001.  
XX  
XX 30-MAR-2001; 2001WO-US08631.  
XX PF  
XX 31-MAR-2000; 2000US-0540217.  
XX PR  
XX 23-AUG-2000; 2000US-0649167.  
XX  
XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;  
XX  
XX WPI; 2001-639362/73.  
DR N-PSDB; AAS92635.

XX New isolated polynucleotide and encoded polypeptides, useful in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits and to assess  
XX biodiversity

XX Claim 20; SEQ ID NO 58807; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and  
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,  
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
XX and gene mapping, and in recombinant production of (II). The  
XX polynucleotides are also used in diagnostics as expressed sequence tags  
XX for identifying expressed genes. (I) is useful in gene therapy techniques  
XX to restore normal activity of (II) or to treat disease states involving  
XX (II). (II) is useful for generating antibodies against it, detecting or  
XX quantitating a polypeptide in tissue, as molecular weight markers and as  
XX a food supplement. (II) and its binding partners are useful in medical

CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 730 AA;  
SQ  
Query Match 50.08; Score 49; DB 22; Length 730;  
Best Local Similarity 43.88; Pred. No. 89;  
Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;  
QY 2 KAPAHLSLPDQALKR 17  
: | | | | | | | | | |  
Db 637 EAPSYIAPDPFVLKQ 652

RESULT 13  
AAB94335  
ID AAB94335 standard; Protein: 324 AA.  
XX  
AC AAB94335;  
XX  
DT 26-JUN-2001 (first entry)  
XX  
DE Human protein sequence SEQ ID NO:14834.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602  
XX full-length cDNAs defined in the specification, and for the detection  
XX and/or diagnosis of the abnormality of the proteins encoded by the  
XX full-length cDNAs

XX Claim 8; SEQ ID 14834; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602  
XX full-length cDNAs defined in the specification. Where a primer set  
XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
XX to the complementary strand of a polynucleotide which comprises one of  
XX the 5602 nucleotide sequences defined in the specification, where the  
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
XX of an oligonucleotide comprising a sequence complementary to the  
XX complementary strand of a polynucleotide which comprises a 5'-end  
XX sequence and an oligonucleotide comprising a sequence complementary to a  
XX polynucleotide which comprises a 3'-end sequence, where the

CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.

XX SQ Sequence 324 AA;  
 Query Match 48.0%; Score 47; DB 22; Length 324;  
 Best Local Similarity 52.9%; Pred. No. 71;  
 Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 EKAPAHLSLPDQALKR 17  
 I: | | | | | | | | | |  
 Db 175 ERPPVHDQLQDPRLLR 191

RESULT 14  
 AAM39674  
 ID AAM39674 standard; Protein; 828 AA.

XX AAM39674;  
 XX 22-OCT-2001 (first entry)  
 XX Human polypeptide SEQ ID NO 2819.  
 XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia.

OS Homo sapiens.  
 XX WO200153312-A1.  
 XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US34263.  
 XX 21-JAN-2000; 2000US-0488725.  
 XX 25-APR-2000; 2000US-0552317.  
 XX 09-JUL-2000; 2000US-0598042.  
 XX 19-JUL-2000; 2000US-0620312.  
 XX 03-AUG-2000; 2000US-0653450.  
 XX 14-SEP-2000; 2000US-0662191.  
 XX 19-OCT-2000; 2000US-0693036.  
 XX 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
 XX WPI; 2001-442253/47.  
 DR N-PSDB; AA158830.

XX Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -  
 XX Example 4; SEQ ID NO 2819; 10078pp; English.  
 XX The invention relates to human nucleic acids (AA157798-AA161369) and

CC the encoded polypeptides (AAM38642-AAH42213) with nootropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.

XX SQ Sequence 828 AA;  
 Query Match 48.0%; Score 47; DB 22; Length 828;  
 Best Local Similarity 52.9%; Pred. No. 2e+02;  
 Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 EKAPAHLSLPDQALKR 17  
 I: | | | | | | | | | |  
 Db 679 ERPPVHDQLQDPRLLR 695

RESULT 15  
 AAB92977  
 ID AAB92977 standard; Protein; 828 AA.

XX AAB92977;  
 XX 26-JUN-2001 (first entry)  
 XX Human protein sequence SEQ ID NO:11690.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.  
 XX Homo sapiens.  
 XX EP1074617-A2.  
 XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.  
 XX 29-JUL-1999; 99JP-0248036.  
 XX 27-AUG-1999; 99JP-0300253.  
 XX 11-JAN-2000; 2000JP-0118776.  
 XX 02-MAY-2000; 2000JP-0183767.  
 XX 09-JUN-2000; 2000JP-0241899.  
 XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -

XX Claim 8; SEQ ID 11690; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesising 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination

CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in  
CC the specification. The primer sets can be used in antisense therapy and  
CC in gene therapy. The primers are useful for synthesising polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to  
CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632  
CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.

XX  
SQ Sequence 828 AA;

Query Match 48.0%; Score 47; DB 22; Length 828;

Best Local Similarity 52.9%; Pred. NO. 2e+02;

Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 EKAPAHLSLPDPQALKR 17

I: | | | | | | | | | |

Db 679 ERPPVHDQLQDPRLKLR 695

Search completed: July 8, 2003, 16:24:16  
Job time : 48.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2003, 16:21:00 ; Search time 11.6923 Seconds  
(without alignments)  
47.812 Million cell updates/sec

Title: US-09-824-647-6

Perfect score: 98

Sequence: 1 EKAPAHLSLPDPQALKRDV 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 2942292 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents.AA.\*

- 1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/1aa/PCTUS.COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	DB ID	Description
1	98	100.0	19	4	US-08-991-862-6	Sequence 6, Appl
2	98	100.0	593	1	US-07-668-648-4	Sequence 4, Appl
3	98	100.0	593	2	US-08-429-998-4	Sequence 4, Appl
4	98	100.0	593	2	US-08-431-333-4	Sequence 4, Appl
5	98	100.0	593	4	US-08-991-862-17	Sequence 17, Appl
6	98	100.0	593	5	PCT-US91-02321-4	Sequence 4, Appl
7	63	64.3	589	1	US-07-668-648-2	Sequence 2, Appl
8	63	64.3	589	2	US-08-429-998-2	Sequence 2, Appl
9	63	64.3	589	2	US-08-431-333-2	Sequence 2, Appl
10	63	64.3	589	5	PCT-US91-02321-2	Sequence 2, Appl
11	53	54.1	179	1	US-07-668-648-8	Sequence 8, Appl
12	53	54.1	179	2	US-08-429-998-8	Sequence 8, Appl
13	53	54.1	179	2	US-08-431-333-8	Sequence 8, Appl
14	53	54.1	179	5	PCT-US91-02321-8	Sequence 8, Appl
15	50	51.0	589	4	US-08-991-862-2	Sequence 2, Appl
16	44	44.9	19	4	US-08-991-862-3	Sequence 3, Appl
17	44	44.9	589	1	US-07-668-648-6	Sequence 6, Appl
18	44	44.9	589	2	US-08-429-998-6	Sequence 6, Appl
19	44	44.9	589	2	US-08-431-333-6	Sequence 6, Appl
20	44	44.9	589	5	PCT-US91-02321-6	Sequence 6, Appl
21	41	41.8	409	3	US-08-807-342B-8	Sequence 8, Appl
22	41	41.8	745	1	US-08-136-277-2	Sequence 2, Appl
23	41	41.8	745	2	US-08-479-403-2	Sequence 2, Appl
24	41	41.8	745	3	US-08-835-734-2	Sequence 2, Appl
25	40.5	41.3	408	4	US-09-315-793-32	Sequence 32, Appl
26	40	40.8	131	2	US-08-525-864A-6	Sequence 6, Appl
27	40	40.8	139	3	US-08-753-007A-33	Sequence 33, Appl

28 40 40.8 139 4 US-09-398-496-33 Sequence 33, Appl  
29 40 40.8 181 3 US-08-753-007A-4 Sequence 4, Appl  
30 40 40.8 181 4 US-09-398-496-4 Sequence 4, Appl  
31 40 40.8 203 1 US-08-233-609-3 Sequence 3, Appl  
32 40 40.8 203 1 US-08-444-083-3 Sequence 3, Appl  
33 40 40.8 203 1 US-08-286-304-3 Sequence 3, Appl  
34 40 40.8 203 1 US-08-442-745-3 Sequence 3, Appl  
35 40 40.8 203 1 US-08-443-129-3 Sequence 3, Appl  
36 40 40.8 203 1 US-08-443-952-3 Sequence 3, Appl  
37 40 40.8 203 1 US-08-443-130-3 Sequence 3, Appl  
38 40 40.8 203 3 US-09-106-182-3 Sequence 3, Appl  
39 40 40.8 203 3 US-08-898-911-3 Sequence 3, Appl  
40 40 40.8 203 5 PCT-US95-04467-3 Sequence 3, Appl  
41 40 40.8 330 2 US-08-525-864A-4 Sequence 4, Appl  
42 40 40.8 407 3 US-08-753-007A-6 Sequence 6, Appl  
43 40 40.8 407 4 US-09-398-496-6 Sequence 6, Appl  
44 40 40.8 469 3 US-08-753-007A-8 Sequence 8, Appl  
45 40 40.8 469 4 US-09-398-496-8 Sequence 8, Appl

ALIGNMENTS

RESULT 1  
US-08-991-862-6  
; Sequence 6, Application US/08991862  
; Patent No. 6309826  
; GENERAL INFORMATION:  
; APPLICANT: Serrero, Ginette  
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
; FILE REFERENCE: 29996.488/P001-A  
; CURRENT APPLICATION NUMBER: US/08/991,862  
; EARLIER FILING DATE: 1998-08-17  
; EARLIER APPLICATION NUMBER: 08/863,862  
; EARLIER FILING DATE: 1997-05-23  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Human granulin  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(19)  
; OTHER INFORMATION: Internal peptide of human GP88 used to develop  
; OTHER INFORMATION: neutralizing anti-human GP88 monoclonal antibody.  
US-08-991-862-6

Query Match 100.0%; Score 98; DB 4; Length 19;  
Best Local Similarity 100.0%; Pred. No. 8.7e-10; Indels 0; Gaps 0;  
Matches 19; Conservative 0; Mismatches 0;  
QY 1 EKAPAHLSLPDPQALKRDV 19  
Db 1 EKAPAHLSLPDPQALKRDV 19  
RESULT 2  
US-07-668-648-4  
; Sequence 4, Application US/07668648  
; Patent No. 5416192  
; GENERAL INFORMATION:  
; APPLICANT: Shoyab, Mohammed  
; APPLICANT: Plowman, Gregory D.  
; TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH  
; TITLE OF INVENTION: MODULATING PROTEINS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA



ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
FILING DATE: 19910819  
APPLICATION NUMBER: US/07/668,648  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Misrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELEPHONE: (212)790-9090  
TELEFAX: (212) 869-9741  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 593 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-668-648-4

Query Match 100.0%; Score 98; DB 1; Length 593;  
Best Local Similarity 100.0%; Pred. No. 4.9e-08;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKAPAHLSLPDPQALKRDV 19  
DB 346 EKAPAHLSLPDPQALKRDV 364

RESULT 3  
US-08-429-998-4  
Sequence 4, Application US/08429998  
Patent No. 5885961  
GENERAL INFORMATION:  
APPLICANT: Shoyab, Mohammed  
TITLE OF INVENTION: EPIITHELINS: NOVEL CYSTEINE-RICH GROWTH  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
FILING DATE: 27-APR-1995  
APPLICATION NUMBER: US/08/429,998  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/668,648  
FILING DATE: 13-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Misrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELEPHONE: (212)790-9090  
TELEFAX: (212) 869-9741  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 593 amino acids

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-429-998-4  
Query Match 100.0%; Score 98; DB 2; Length 593;  
Best Local Similarity 100.0%; Pred. No. 4.9e-08;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKAPAHLSLPDPQALKRDV 19  
DB 346 EKAPAHLSLPDPQALKRDV 364

RESULT 4  
US-08-431-333-4  
Sequence 4, Application US/08431333  
Patent No. 5965723  
GENERAL INFORMATION:  
APPLICANT: Shoyab, Mohammed  
TITLE OF INVENTION: EPIITHELINS: NOVEL CYSTEINE-RICH GROWTH  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
FILING DATE: 27-APR-1995  
APPLICATION NUMBER: US/08/431,333  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/668,648  
FILING DATE: 13-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Misrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELEPHONE: (212)790-9090  
TELEFAX: (212) 869-9741  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 593 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-431-333-4

Query Match 100.0%; Score 98; DB 2; Length 593;  
Best Local Similarity 100.0%; Pred. No. 4.9e-08;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKAPAHLSLPDPQALKRDV 19  
DB 346 EKAPAHLSLPDPQALKRDV 364

RESULT 5  
US-08-991-862-17  
Sequence 17, Application US/08991862  
Patent No. 6309826  
GENERAL INFORMATION:  
APPLICANT: Serrero, Ginette

;; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
;; FILE REFERENCE: 29996.488/P001-A  
;; CURRENT APPLICATION NUMBER: US/08/991.862  
;; CURRENT FILING DATE: 1998-08-17  
;; EARLIER APPLICATION NUMBER: 08/863.862  
;; EARLIER FILING DATE: 1997-05-23  
;; NUMBER OF SEQ ID NOS: 17  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 17  
;; LENGTH: 593  
;; TYPE: PRT  
;; ORGANISM: Human GP88 cDNA  
US-08-991-862-17

Query Match 100.0%; Score 98; DB 4; Length 593;  
Best Local Similarity 100.0%; Pred. No. 4.9e-08;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKAPAHLSLPDPQALKRDV 19  
Db 346 EKAPAHLSLPDPQALKRDV 364

RESULT 6  
PCT-US91-02321-4  
;; Sequence 4, Application PCT/US9102321  
;; GENERAL INFORMATION:  
;; APPLICANT: Shoyab, Mohammed  
;; APPLICANT: Plovman, Gregory D.  
;; TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH  
;; TITLE OF INVENTION: MODULATING PROTEINS  
;; NUMBER OF SEQUENCES: 12  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Bristol-Myers Squibb Company  
;; STREET: 3005 First Avenue  
;; CITY: Seattle  
;; STATE: Washington  
;; COUNTRY: USA  
;; ZIP: 98121  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; FILING DATE: 19910403  
;; APPLICATION NUMBER: PCT/US91/02321  
;; CLASSIFICATION: 514  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Poor, Brian W.  
;; REGISTRATION NUMBER: 32,928  
;; REFERENCE/DOCKET NUMBER: 080071A-PC  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (206)728-4800  
;; TELEFAX: (206)448-4775  
;; INFORMATION FOR SEQ ID NO: 4:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 593 amino acids  
;; TYPE: AMINO ACID  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
PCT-US91-02321-4

Query Match 100.0%; Score 98; DB 5; Length 593;  
Best Local Similarity 100.0%; Pred. No. 4.9e-08;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKAPAHLSLPDPQALKRDV 19  
Db 346 EKAPAHLSLPDPQALKRDV 364

RESULT 7

US-07-668-648-2  
;; Sequence 2, Application US/07668648  
;; Patent No. 5416192  
;; GENERAL INFORMATION:  
;; APPLICANT: Shoyab, Mohammed  
;; APPLICANT: Plovman, Gregory D.  
;; TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH  
;; TITLE OF INVENTION: MODULATING PROTEINS  
;; NUMBER OF SEQUENCES: 12  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Pennie & Edmonds  
;; STREET: 1155 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/668.648  
;; FILING DATE: 19910819  
;; CLASSIFICATION: 514  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Mistrock, S. Leslie  
;; REGISTRATION NUMBER: 18,872  
;; REFERENCE/DOCKET NUMBER: 5624-161-999  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212)790-9090  
;; TELEFAX: (212) 869-9741  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 589 amino acids  
;; TYPE: AMINO ACID  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-07-668-648-2

Query Match 64.3%; Score 63; DB 1; Length 589;  
Best Local Similarity 68.4%; Pred. No. 0.024;  
Matches 13; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 EKAPAHLSLPDPQALKRDV 19  
Db 344 KKVTAASLSPDPQILKNDV 362

RESULT 8  
US-08-429-998-2  
;; Sequence 2, Application US/08429998  
;; Patent No. 5885961  
;; GENERAL INFORMATION:  
;; APPLICANT: Shoyab, Mohammed  
;; APPLICANT: Plovman, Gregory D.  
;; TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH  
;; TITLE OF INVENTION: MODULATING PROTEINS  
;; NUMBER OF SEQUENCES: 12  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Pennie & Edmonds  
;; STREET: 1155 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/429,998

US-08-431-333-2

ADDRESS: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York

US-08-431-333-2

; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/668,648  
; FILING DATE: 19910819  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mistrock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 5624-161-999  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-9741  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 179 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-07-668-648-8

Query Match 54.1%; Score 53; DB 1; Length 179;  
Best Local Similarity 63.2%; Pred. No. 0.25;  
Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 EKAPAHLSLPDPOALKRDV 19  
:||||||| | | :||  
Db 118 KKAPAHLSLLDLGAVEGDV 136

RESULT 12  
US-08-429-998-8  
; Sequence 8, Application US/08429998  
; Patent No. 5885961  
; GENERAL INFORMATION:  
; APPLICANT: Shoyab, Mohammed  
; TITLE OF INVENTION: EPIITHELINS: NOVEL CYSTEINE-RICH GROWTH  
; TITLE OF INVENTION: MODULATING PROTEINS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/429,998  
; FILING DATE: 27-APR-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION NUMBER:  
; FILING DATE: 13-MAR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mistrock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 5624-161-999  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-9741  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:

; LENGTH: 179 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-429-998-8

Query Match 54.1%; Score 53; DB 2; Length 179;  
Best Local Similarity 63.2%; Pred. No. 0.25;  
Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 EKAPAHLSLPDPOALKRDV 19  
:||||||| | | :||  
Db 118 KKAPAHLSLLDLGAVEGDV 136

RESULT 13  
US-08-431-333-8  
; Sequence 8, Application US/08431333  
; Patent No. 5965723  
; GENERAL INFORMATION:  
; APPLICANT: Shoyab, Mohammed  
; APPLICANT: Plowman, Gregory D.  
; TITLE OF INVENTION: EPIITHELINS: NOVEL CYSTEINE-RICH GROWTH  
; TITLE OF INVENTION: MODULATING PROTEINS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/431,333  
; FILING DATE: 27-APR-1995  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/668,648  
; FILING DATE: 13-MAR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mistrock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 5624-161-999  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-9741  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 179 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-431-333-8

Query Match 54.1%; Score 53; DB 2; Length 179;  
Best Local Similarity 63.2%; Pred. No. 0.25;  
Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 EKAPAHLSLPDPOALKRDV 19  
:||||||| | | :||  
Db 118 KKAPAHLSLLDLGAVEGDV 136

RESULT 14  
PCT-US91-02321-8  
; Sequence 8, Application PC/TUS9102321  
; GENERAL INFORMATION:  
; APPLICANT: Shoyab, Mohammed

APPLICANT: Plowman, Gregory D.  
TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH  
MODULATING PROTEINS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Bristol-Myers Squibb Company  
STREET: 3005 First Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98121  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/02321  
FILING DATE: 19910403  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Poor, Brian W.  
REGISTRATION NUMBER: 32,928  
REFERENCE/DOCKET NUMBER: ON0071A-PC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)728-4800  
TELEFAX: (206)448-4775  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 179 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US91-02321-8

Query Match 54.1%; Score 53; DB 5; Length 179;  
Best Local Similarity 63.2%; Pred. No. 0.25;  
Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 EKAPAHLSLPDPQALKRDV 19  
:||||||| | | :||  
Db 118 KKAPAHLSLIDLGAVEGDV 136

RESULT 15  
US-08-991-862-2  
; Sequence 2, Application US/08991862  
; Patent No. 6309826  
; GENERAL INFORMATION:  
; APPLICANT: Sertero, Ginette  
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
; FILE REFERENCE: 29996.488/P001-A  
; CURRENT APPLICATION NUMBER: US/08/991.862  
; CURRENT FILING DATE: 1998-08-17  
; EARLIER APPLICATION NUMBER: 08/863.862  
; EARLIER FILING DATE: 1997-05-23  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 589  
; TYPE: PRT  
; ORGANISM: Mouse epithelin/granulin  
US-08-991-862-2

Query Match 51.0%; Score 50; DB 4; Length 589;  
Best Local Similarity 61.1%; Pred. No. 3.1;  
Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 EKAPAHLSLPDPQALKRD 18  
:| | | | | | | |  
Db 344 KVIAPLRLPDPQILKSD 361

Search completed: July 8, 2003, 16:30:20  
Job time : 12.6923 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2003, 16:24:26 ; Search time 18.5128 Seconds  
(without alignments)  
119.483 Million cell updates/sec

Title: US-09-824-647-6  
Perfect score: 98  
Sequence: 1 EKAPAHLSLPDPQALKRDV 19

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 116419773 residues

Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	98	100.0	19	9	US-09-824-647-6
2	98	100.0	19	9	US-10-218-509-6
3	98	100.0	19	9	US-10-281-160-6
4	98	100.0	19	10	US-09-813-156-6
5	98	100.0	19	10	US-09-824-807-6
6	98	100.0	593	9	US-09-824-647-17
7	98	100.0	593	9	US-10-218-509-17
8	98	100.0	593	9	US-10-281-160-17
9	98	100.0	593	10	US-09-813-156-17
10	98	100.0	593	10	US-09-824-807-17
11	98	100.0	621	10	US-09-925-301-1416
12	50	51.0	589	9	US-09-824-647-2
13	50	51.0	589	9	US-10-218-509-2
14	50	51.0	589	9	US-10-281-160-2
15	50	51.0	589	10	US-09-813-156-2
16	50	51.0	589	10	US-09-824-807-2
17	48	49.0	712	9	US-09-364-847-51
18	46	46.9	2115	9	US-10-208-948-4
19	46	46.9	2150	9	US-10-208-948-27

Sequence 34773, A  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 5211, Ap  
Sequence 8875, Ap  
Sequence 22, Appl  
Sequence 62, Appl  
Sequence 3343, Ap  
Sequence 8343, Ap  
Sequence 2283, Ap  
Sequence 2283, Ap  
Sequence 9516, Ap  
Sequence 12577, A  
Sequence 4349, Ap  
Sequence 64, Appl  
Sequence 3, Appli  
Sequence 7738, Ap  
Sequence 54, Appl  
Sequence 33516, A  
Sequence 43750, A  
Sequence 145, App  
Sequence 305, App  
Sequence 33, Appl

20 45 45.9 59 10 US-09-864-761-34773  
21 44 44.9 19 9 US-09-824-647-3  
22 44 44.9 19 9 US-10-218-509-3  
23 44 44.9 19 9 US-10-281-160-3  
24 44 44.9 19 10 US-09-813-156-3  
25 44 44.9 19 10 US-09-824-807-3  
26 43.5 44.4 619 10 US-09-815-242-5211  
27 43 43.9 332 9 US-10-156-761-8875  
28 43 43.9 479 9 US-09-972-288-22  
29 43 43.9 479 9 US-10-161-572-62  
30 43 43.9 656 9 US-10-128-714-3343  
31 43 43.9 745 9 US-10-128-714-8343  
32 42 42.9 68 9 US-09-796-692-2283  
33 42 42.9 68 9 US-10-040-862-2283  
34 42 42.9 962 9 US-10-156-761-9516  
35 42 42.9 1225 9 US-10-156-761-12577  
36 41 41.8 168 9 US-09-738-626-4349  
37 41 41.8 276 9 US-09-870-759-64  
38 41 41.8 276 9 US-10-043-649-3  
39 41 41.8 294 9 US-10-156-761-7738  
40 41 41.8 410 9 US-10-038-010-54  
41 40.5 41.3 134 10 US-09-864-761-33516  
42 40 40.8 53 10 US-09-864-761-43750  
43 40 40.8 86 9 US-10-097-065-145  
44 40 40.8 96 10 US-09-912-020-305  
45 40 40.8 139 12 US-10-096-241-33

## ALIGNMENTS

RESULT 1  
US-09-824-647-6  
; Sequence 6, Application US/09824647  
; Publication No. US20020183270A1  
; GENERAL INFORMATION:  
; APPLICANT: Serrero, GINETTE  
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
; FILE REFERENCE: Z9996.488/P001-A  
; CURRENT APPLICATION NUMBER: US/09/824,647  
; CURRENT FILING DATE: 2001-04-04  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/991,862  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/863,862  
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-05-23  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 19  
; TYPE: PPT  
; ORGANISM: Human granulin  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(19)  
; OTHER INFORMATION: Internal peptide of human GP88 used to develop  
; OTHER INFORMATION: neutralizing anti-human GP88 monoclonal antibody.  
US-09-824-647-6

Query Match 100.0%; Score 98; DB 9; Length 19;  
Best Local Similarity 100.0%; Pred. No. 7.8e-09;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EKAPAHLSLPDPQALKRDV 19  
|||||  
Db 1 EKAPAHLSLPDPQALKRDV 19

RESULT 2  
US-10-218-509-6  
; Sequence 6, Application US/10218509  
; Publication No. US20030092661A1  
; GENERAL INFORMATION:  
; APPLICANT: Serrero, GINETTE

;; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
;; FILE REFERENCE: Z9996.488/P001-A  
;; CURRENT APPLICATION NUMBER: US/10/218,509  
;; CURRENT FILING DATE: 2002-08-15  
;; PRIOR APPLICATION NUMBER: 08/991,862  
;; PRIOR FILING DATE: 1998-08-17  
;; PRIOR APPLICATION NUMBER: 08/863,862  
;; PRIOR FILING DATE: 1997-05-23  
;; NUMBER OF SEQ ID NOS: 17  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 6  
;; LENGTH: 19  
;; TYPE: PRT  
;; ORGANISM: Human granulin  
;; FEATURE:  
;; NAME/KEY: PEPTIDE  
;; LOCATION: (1)..(19)  
;; OTHER INFORMATION: Internal peptide of human GP88 used to develop  
;; OTHER INFORMATION: neutralizing anti-human GP88 monoclonal antibody.  
US-10-218-509-6

Query Match 100.0%; Score 98; DB 9; Length 19;  
Best Local Similarity 100.0%; Pred. No. 7.8e-09;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKAPAHLSLPDPQALKRDV 19  
|||  
DB 1 EKAPAHLSLPDPQALKRDV 19

RESULT 3  
US-10-281-160-6  
;; Sequence 6, Application US/10281160  
;; Publication No. US20030108950A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Serriero, Ginette  
;; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
;; FILE REFERENCE: Z9996.488/P001-A  
;; CURRENT APPLICATION NUMBER: US/10/281,160  
;; CURRENT FILING DATE: 2002-10-28  
;; PRIOR APPLICATION NUMBER: 08/991,862  
;; PRIOR FILING DATE: 1998-08-17  
;; PRIOR APPLICATION NUMBER: 08/863,862  
;; PRIOR FILING DATE: 1997-05-23  
;; NUMBER OF SEQ ID NOS: 17  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 6  
;; LENGTH: 19  
;; TYPE: PRT  
;; ORGANISM: Human granulin  
;; FEATURE:  
;; NAME/KEY: PEPTIDE  
;; LOCATION: (1)..(19)  
;; OTHER INFORMATION: Internal peptide of human GP88 used to develop  
;; OTHER INFORMATION: neutralizing anti-human GP88 monoclonal antibody.  
US-10-281-160-6

Query Match 100.0%; Score 98; DB 9; Length 19;  
Best Local Similarity 100.0%; Pred. No. 7.8e-09;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKAPAHLSLPDPQALKRDV 19  
|||  
DB 1 EKAPAHLSLPDPQALKRDV 19

RESULT 4  
US-09-813-156-6  
;; Sequence 6, Application US/09813156  
;; Patent No. US20020061859A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Serriero, Ginette  
;; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS

;; FILE REFERENCE: Z9996.488/P001-A  
;; CURRENT APPLICATION NUMBER: US/09/813,156  
;; CURRENT FILING DATE: 2001-03-21  
;; PRIOR APPLICATION NUMBER: 08/991,862  
;; PRIOR FILING DATE: 1997-12-16  
;; PRIOR APPLICATION NUMBER: 08/863,862  
;; PRIOR FILING DATE: 1997-05-23  
;; NUMBER OF SEQ ID NOS: 17  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 6  
;; LENGTH: 19  
;; TYPE: PRT  
;; ORGANISM: Human granulin  
;; FEATURE:  
;; NAME/KEY: PEPTIDE  
;; LOCATION: (1)..(19)  
;; OTHER INFORMATION: Internal peptide of human GP88 used to develop  
;; OTHER INFORMATION: neutralizing anti-human GP88 monoclonal antibody.  
US-09-813-156-6

Query Match 100.0%; Score 98; DB 10; Length 19;  
Best Local Similarity 100.0%; Pred. No. 7.8e-09;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKAPAHLSLPDPQALKRDV 19  
|||  
DB 1 EKAPAHLSLPDPQALKRDV 19

RESULT 5  
US-09-824-807-6  
;; Sequence 6, Application US/09824807  
;; Patent No. US20020094966A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Serriero, Ginette  
;; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
;; FILE REFERENCE: Z9996.488/P001-A  
;; CURRENT APPLICATION NUMBER: US/09/824,807  
;; CURRENT FILING DATE: 2001-04-04  
;; PRIOR APPLICATION NUMBER: 08/991,862  
;; PRIOR FILING DATE: 1997-12-16  
;; PRIOR APPLICATION NUMBER: 08/863,862  
;; PRIOR FILING DATE: 1997-05-23  
;; NUMBER OF SEQ ID NOS: 17  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 6  
;; LENGTH: 19  
;; TYPE: PRT  
;; ORGANISM: Human granulin  
;; FEATURE:  
;; NAME/KEY: PEPTIDE  
;; LOCATION: (1)..(19)  
;; OTHER INFORMATION: Internal peptide of human GP88 used to develop  
;; OTHER INFORMATION: neutralizing anti-human GP88 monoclonal antibody.  
US-09-824-807-6

Query Match 100.0%; Score 98; DB 10; Length 19;  
Best Local Similarity 100.0%; Pred. No. 7.8e-09;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKAPAHLSLPDPQALKRDV 19  
|||  
DB 1 EKAPAHLSLPDPQALKRDV 19

RESULT 6  
US-09-824-647-17  
;; Sequence 17, Application US/09824647  
;; Publication No. US20020183270A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Serriero, Ginette  
;; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
;; FILE REFERENCE: Z9996.488/P001-A

```
; CURRENT APPLICATION NUMBER: US/09/824.647
; CURRENT FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/991.862
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/863.862
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-05-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Human GP88 cdna
US-09-824-647-17

Query Match      100.0%; Score 98; DB 9; Length 593;
Best Local Similarity 100.0%; Pred. No. 3.8e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 EKAPAHLSLPDPQALKRDV 19
Db      346 EKAPAHLSLPDPQALKRDV 364
|||||

RESULT 7
US-10-218-509-17
; Sequence 17, Application US/10218509
; Publication No. US20030092661A1
; GENERAL INFORMATION:
; APPLICANT: Serrero, Ginette
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
; FILE REFERENCE: Z9996.488/P001-A
; CURRENT APPLICATION NUMBER: US/10/218.509
; CURRENT FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: 08/991.862
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 08/863.862
; PRIOR FILING DATE: 1997-05-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Human GP88 cdna
US-10-218-509-17

Query Match      100.0%; Score 98; DB 9; Length 593;
Best Local Similarity 100.0%; Pred. No. 3.8e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 EKAPAHLSLPDPQALKRDV 19
Db      346 EKAPAHLSLPDPQALKRDV 364
|||||

RESULT 8
US-10-281-160-17
; Sequence 17, Application US/10281160
; Publication No. US20030108950A1
; GENERAL INFORMATION:
; APPLICANT: Serrero, Ginette
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
; FILE REFERENCE: Z9996.488/P001-A
; CURRENT APPLICATION NUMBER: US/10/281.160
; CURRENT FILING DATE: 2002-10-28
; PRIOR APPLICATION NUMBER: US/08/991.862
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 08/863.862
; PRIOR FILING DATE: 1997-05-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Human GP88 cdna
US-10-281-160-17

Query Match      100.0%; Score 98; DB 9; Length 593;
Best Local Similarity 100.0%; Pred. No. 3.8e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 EKAPAHLSLPDPQALKRDV 19
Db      346 EKAPAHLSLPDPQALKRDV 364
|||||

RESULT 9
US-09-813-156-17
; Sequence 17, Application US/09813156
; Patent No. US20020061859A1
; GENERAL INFORMATION:
; APPLICANT: Serrero, Ginette
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
; FILE REFERENCE: Z9996.488/P001-A
; CURRENT APPLICATION NUMBER: US/09/813.156
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 08/991.862
; PRIOR FILING DATE: 1997-12-16
; PRIOR APPLICATION NUMBER: 08/863.862
; PRIOR FILING DATE: 1997-05-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Human GP88 cdna
US-09-813-156-17

Query Match      100.0%; Score 98; DB 10; Length 593;
Best Local Similarity 100.0%; Pred. No. 3.8e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 EKAPAHLSLPDPQALKRDV 19
Db      346 EKAPAHLSLPDPQALKRDV 364
|||||

RESULT 10
US-09-824-807-17
; Sequence 17, Application US/09824807
; Patent No. US20020094966A1
; GENERAL INFORMATION:
; APPLICANT: Serrero, Ginette
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
; FILE REFERENCE: Z9996.488/P001-A
; CURRENT APPLICATION NUMBER: US/09/824.807
; CURRENT FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 08/991.862
; PRIOR FILING DATE: 1997-12-16
; PRIOR APPLICATION NUMBER: 08/863.862
; PRIOR FILING DATE: 1997-05-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Human GP88 cdna
US-09-824-807-17

Query Match      100.0%; Score 98; DB 10; Length 593;
Best Local Similarity 100.0%; Pred. No. 3.8e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 EKAPAHLSLPDPQALKRDV 19
Db      346 EKAPAHLSLPDPQALKRDV 364
|||||
```



```

; APPLICANT: Serrero, Ginette
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
; FILE REFERENCE: Z9996.488/P001-A
; CURRENT APPLICATION NUMBER: US/10/218,509
; CURRENT FILING DATE: 2002-08-15

```

```

; SEQ ID NO 2
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Mouse epithelin/granulin
US-09-813-156-2

```

Query Match 51.0%; Score 50; DB 10; Length 589;  
Best Local Similarity 61.1%; Pred. No. 10;  
Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;  
QY 1 EKAPAHLSLPDPQALKRD 18  
Db : | | | | | | | | | |  
344 KKVIAPLRLPDPQILKSD 361

Search completed: July 8, 2003, 16:31:42  
Job time : 19.5128 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2003, 16:19:45 ; Search time 19.2436 Seconds  
(without alignments)  
94.918 Million cell updates/sec

Title: US-09-824-647-6

Perfect score: 98

Sequence: 1 EKAPAHLSLPDQALRKDV 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_73.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	98	100.0	593	1 GYHU	granulin precursor
2	63	64.3	589	2 B38128	epithelin/granulin
3	46	46.9	1857	2 T50513	hypothetical prote
4	46	46.9	2143	2 G96595	hypothetical prote
5	45	45.9	374	2 S76322	hypothetical prote
6	44.5	45.4	560	2 I50372	hypothetical prote
7	44	44.9	187	2 H72559	hypothetical prote
8	44	44.9	223	2 T22911	hypothetical prote
9	44	44.9	248	2 H90221	ribonuclease PH (r
10	44	44.9	464	2 H96700	protein F12A21.5 [
11	44	44.9	589	2 C38128	epithelin/granulin
12	44	44.9	722	2 G75563	phosphate acetyltr
13	43.5	44.4	619	2 F82984	hypothetical prote
14	43	43.9	366	2 AH2237	septum site-determ
15	43	43.9	406	2 C98263	hypothetical prote
16	43	43.9	406	2 AG3021	murein hydrolase [
17	43	43.9	478	2 I33960	PRR2 alpha - human
18	43	43.9	990	2 B49351	bacteriophage N4 a
19	42	42.9	129	2 B83053	secretion protein
20	42	42.9	488	2 T46668	probable 2,3-dehyd
21	42	42.9	701	2 T20892	hypothetical prote
22	42	42.9	1148	2 T18770	probable calcium c
23	42	42.9	1956	2 T00051	hypothetical prote
24	41	41.8	393	2 S64303	cysteine synthase
25	41	41.8	410	2 A59279	beta-arrestin 2 -
26	41	41.8	438	1 ERADC2	early E1B 49k prot
27	41	41.8	558	2 G87414	hypothetical prote
28	41	41.8	710	1 C43497	early transcriptio
29	41	41.8	710	1 I42517	early transcriptio

#### ALIGNMENTS

##### RESULT 1

GYHU

granulin precursor [validated] - human

N:Alternate names: epithelin

N:Contains: granulin A; granulin B; granulin C; granulin D; granulin E; granulin

C:Species: Homo sapiens (man)

C:Date: 30-Sep-1992 #sequence-revision 03-May-1996 #text-change 08-Dec-2000

C:Accession: J01284; A38128; A38118; A36698; B36698; C36698; D36698; A56873

R:Bhandari, V.; Bateman, A.

Biochem. Biophys. Res. Commun. 188, 57-63, 1992

A:Title: Structure and chromosomal location of the human granulin gene.

A:Reference number: J01284; MUID:93038704; PMID:1417868

A:Accession: J01284

A:Molecule type: DNA

A:Residues: 1-593 <BHA>

R:Plowman, G.D.; Green, J.M.; Neubauer, M.G.; Buckley, S.D.; McDonald, V.L.; Toda

J. Biol. Chem. 267, 13073-13078, 1992

A:Title: The epithelin precursor encodes two proteins with opposing activities or

A:Reference number: A38128; MUID:92317004; PMID:1618805

A:Accession: A38128

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-593 <PIQ>

A:Cross-references: GB:X62320; NID:931192; PIDN:CAA44196.1; PID:931193

R:Bhandari, V.; Palfrey, R.G.E.; Bateman, A.

Proc. Natl. Acad. Sci. U.S.A. 89, 1715-1719, 1992

A:Title: Isolation and sequence of the granulin precursor cDNA from human bone m

A:Reference number: A38118; MUID:92179253; PMID:1542665

A:Accession: A38118

A:Molecule type: mRNA

A:Residues: 1-406, 'R', 408-433, 'G', 435-453, 'G', 455-459, 'Q', 461-546, 'A', 548-566, 'R'

A:Cross-references: GB:M75161; NID:9183612; PIDN:AAA58617.1; PID:9183613

A:Note: this sequence has been revised in reference J01284

R:Bateman, A.; Belcourt, D.; Bennett, H.; Lazure, C.; Solomon, S.

Biochem. Biophys. Res. Commun. 173, 1161-1168, 1990

A:Title: Granulins, a novel class of peptide from leukocytes.

A:Reference number: A36698; MUID:91097544; PMID:2268320

A:Accession: A36698

A:Molecule type: protein

A:Residues: 281-336 <BAT>

A:Note: this protein was purified and characterized as granulin A

A:Accession: B36698

A:Molecule type: protein

A:Residues: 206-218, 'H', 220-233 <BA2>

A:Note: this protein was purified and characterized as granulin B

A:Accession: C36698

A:Molecule type: protein

A:Residues: 364-367, 'X', 369-385, 'H', 387-396 <BA3>

A:Note: this protein was purified and characterized as granulin C

A:Accession: D36698

A:Molecule type: protein

A:Residues: 442-446, 'ADTSS', 456-458, 'DG' <BA4>

A:Note: this protein was purified and characterized as granulin D

R.Kardana, A.; Bagshawe, K.D.; Coles, B.; Read, D.; Taylor, M.  
Br. J. Cancer 67, 686-692, 1993

A>Title: Characterisation of UGP and its relationship with beta-core fragment.  
A:Reference number: A56873; MUID:93229246; PMID:8471426

A:Accession: A56873

A:Molecule type: protein

A:Residues: 281-283, 'X', 285-289, 'S', 291-295 <KAR>

A:Experimental source: urine

A>Note: sequence extracted from NCBI backbone (NCBIP:129524)

C:Genetics:

A:Gene: GDB:GRN

A:Cross-references: GDB:136006; OMIM:138945

A:Map position: 17pter-17qter

A:Introns: 46/3; 88/3; 117/1; 154/3; 200/1; 236/3; 279/1; 311/3; 393/3; 471/3; 548/3

C:Superfamily: granulin

C:Keywords: glycoprotein; tandem repeat

F:1-17/Domain: signal sequence #status predicted <SIG>

F:18-593/Product: granulin #status predicted <MAT>

F:18-593/Product: granulin #status predicted <PRO>

F:18-44/Product: paraganulin #status experimental <PGR>

F:18-113/Product: granulin G #status predicted <GRG>

F:123-179/Product: granulin F #status predicted <GRF>

F:206-261/Product: granulin B #status predicted <GRB>

F:281-336/Product: granulin A #status experimental <GRA>

F:364-417/Product: granulin C #status experimental <GRC>

F:442-496/Product: granulin D #status predicted <GRD>

F:518-573/Product: granulin E #status predicted <GRE>

F:368/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 100.0%; Score 98; DB 1; Length 593;

Best Local Similarity 100.0%; Pred. No. 1.1e-07;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKAPAHLSLPDPQALKRDV 19

Db 346 EKAPAHLSLPDPQALKRDV 364

#### RESULT 2

B38128

epithelin/granulin precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 10-Jul-1992 #sequence\_revision 10-Jul-1992 #text\_change 20-Aug-1999

C:Accession: B38128; A36199; B36199; E36698; I53272

R:Plowman, G.D.; Green, J.M.; Neubauber, M.G.; Buckley, S.D.; McDonald, V.L.; Todaro, G.J.

J. Biol. Chem. 267, 13073-13078, 1992

A>Title: The epithelin precursor encodes two proteins with opposing activities on epithe

A:Reference number: A36128; MUID:92317004; PMID:1618805

A:Accession: B38128

A:Molecule type: mRNA

A:Residues: 1-589 <PLO>

A:Cross-references: GB:X62322; NID:956108; PIDN:CAAA4198.1; PID:956109

R:Shoyab, M.; McDonald, V.L.; Byles, C.; Todaro, G.J.; Plowman, G.D.

Proc. Natl. Acad. Sci. U.S.A. 87, 7912-7916, 1990

A>Title: Epithelins 1 and 2: isolation and characterization of two cysteine-rich growth-

A:Reference number: A36199; MUID:91045907; PMID:2236009

A:Accession: A36199

A:Molecule type: protein

A:Residues: 280-300 <SHO>

A:Accession: B36199

A:Molecule type: protein

A:Residues: 205-226 <SH2>

R:Bateman, A.; Belcourt, D.; Bennett, H.; Lazure, C.; Solomon, S.

Biochem. Biophys. Res. Commun. 173, 1161-1168, 1990

A>Title: Granulins, a novel class of peptide from leukocytes.

A:Reference number: A36698; MUID:91097544; PMID:2268320

A:Accession: E36698

A:Molecule type: protein

A:Residues: 279-307, 'SB', 310-324, 'T', 326, 'X', 328, 'Q' <BAT>

R:Bhandari, V.; Giald, A.; Bateman, A.

Endocrinology 133, 2682-2689, 1993

A>Title: The complementary deoxyribonucleic acid sequence, tissue distribution, and cell

A:Reference number: I53272; MUID:94062640; PMID:8243292

A:Accession: I53272

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-200, 'S', 203-388, 'M', 390-589 <RES>

A:Cross-references: GB:M97750; NID:9204223; PIDN:AAAL6903.1; PID:9204224

C:Superfamily: granulin

Query Match 64.3%; Score 63; DB 2; Length 589; #

Best Local Similarity 68.4%; Pred. No. 0.044;

Matches 13; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 EKAPAHLSLPDPQALKRDV 19

Db 344 KKVTAASLSLPDQILKNDV 362

#### RESULT 3

T50513

hypothetical protein T27115\_10 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 21-Jul-2000

C:Accession: T50513

R:Choisme, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.

submitted to the Protein Sequence database, May 2000

A:Reference number: T25102

A:Accession: T50513

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1857 <CHO>

A:Cross-references: EMBL:AL358732

A:Experimental source: cultivar Columbia; BAC clone T27115

C:Genetics:

A:Map position: 3

A:Introns: 295/1; 320/1; 550/3; 567/1; 721/3; 766/3; 801/3; 839/3; 865/2; 892/3; 1

A>Note: T27115\_10

Query Match 46.9%; Score 46; DB 2; Length 1857;

Best Local Similarity 52.9%; Pred. No. 83;

Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 KAPAHLSLPDPQALKRD 18

Db 1336 EAPAHPPAPPPLRRD 1352

#### RESULT 4

G96595

hypothetical protein F7A10.23 [Imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001

C:Accession: G96595

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; A.

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar,

Nature, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; K.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marz

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tai

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: G96595

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-2143 <STO>

A:Cross-references: GB:AE005173; NID:gi0645420; PIDN:AAG21537.1; GSPDB:GN00141

C:Genetics:

A:Gene: F7A10.23

A:Map position: 1

Query Match 46.9%; Score 46; DB 2; Length 2143;

Best Local Similarity 61.5%; Pred. No. 97;

Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 HLSLDPQALKRDV 18  
 DB 372 HLSITDPSAARD 384

# RESULT 5

S76322  
 hypothetical protein - Synechocystis sp. (strain PCC 6803)  
 C:Species: Synechocystis sp.  
 A:Variety: PCC 6803  
 C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 20-Jun-2000  
 C:Accession: S76322  
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda  
 DNA Res. 3, 109-136, 1996  
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis  
 S.  
 A:Reference number: S74322; MUID:97061201; PMID:8905231  
 A:Accession: S76322  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-374 <KAN>  
 A:Cross-references: EMBL:D64000; GB:AB001339; NID:g1001484; PIDN:BAAL0174.1; PID:g100154  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 45.9%; Score 45; DB 2; Length 374;

Best Local Similarity 50.0%; Pred. No. 20;

Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 KAPAHLSLPDQAL 15  
 DB 339 RPPVHLDPDPSKM 352

# RESULT 6

I50372  
 hypothetical protein 2 delta-EF1 region - chicken  
 C:Species: Gallus gallus (chicken)  
 C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 05-May-2000  
 C:Accession: I50372  
 R:Funahashi, J.; Sekido, R.; Murai, K.; Kanachi, Y.; Kondoh, H.  
 Development 119, 433-446, 1993  
 A:Title: Delta-crystallin enhancer binding protein delta EF1 is a zinc finger-homeodomain  
 A:Reference number: I50222; MUID:94116444; PMID:7904558  
 A:Accession: I50372  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-560 <FUN>  
 A:Cross-references: GB:D14316; NID:g391639; PIDN:BAAO3262.1; PID:g391640  
 C:Superfamily: CHD-1 protein; chromobox homology

Query Match 45.4%; Score 44.5; DB 2; Length 560;

Best Local Similarity 55.0%; Pred. No. 38;

Matches 11; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 1 EKAPAHLSLP-DPQALKRDV 19  
 DB 468 ELAPLHKSPDPEERKRV 487

# RESULT 7

H72559  
 hypothetical protein APE1762 - Aeropyrum pernix (strain K1)  
 C:Species: Aeropyrum pernix  
 C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jun-2000  
 C:Accession: H72559  
 R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah  
 awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K  
 DNA Res. 6, 83-101, 1999  
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr  
 A:Reference number: A74450; MUID:99310339; PMID:10382966

A:Accession: H72559

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-187 <KAW>

A:Cross-references: DDBJ:AP000062; NID:g5105244; PIDN:BAAB0765.1; PID:d1044551; I

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE1762

C:Superfamily: Aeropyrum pernix hypothetical protein APE1762

Query Match 44.9%; Score 44; DB 2; Length 187;

Best Local Similarity 71.4%; Pred. No. 13;

Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 HLSLDPQALKRDV 19

DB 90 HRSPLHPGALKRVV 103

# RESULT 8

T22911

hypothetical protein F58E10.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T22911

R:McMurray, A.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19637

A:Accession: T22911

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-223 <WIL>

A:Cross-references: EMBL:Z81555; PIDN:CAB04512.1; GSPDB:GN00023; CESP:F58E10.6

A:Experimental source: clone F58E10

C:Genetics:

A:Gene: CESP:F58E10.6

A:Map position: 5

A:Introns: 64/3; 166/1

Query Match 44.9%; Score 44; DB 2; Length 223;

Best Local Similarity 50.0%; Pred. No. 16;

Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 4 PAHLSLPDQALKRDV 19

DB 47 PAYLKPENPEVAKQYV 62

# RESULT 9

H90221

ribonuclease PH (rph) [imported] - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus

C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 15-Jun-2001

C:Accession: H90221

R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J

Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; R

arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A:Description: Sulfolobus solfataricus complete genome.

A:Reference number: A99139

A:Accession: H90221

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-248 <KUR>

A:Cross-references: GB:AE006641; NID:g13813903; PIDN:AAK41031.1; GSPDB:GN00155

C:Genetics:

A:Gene: rph

C:Superfamily: tRNA nucleotidyltransferase

Query Match 44.9%; Score 44; DB 2; Length 248;

Best Local Similarity 56.3%; Pred. No. 18;

Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 EKAPHLSPDQALK 16  
I I I I I I I I  
Db 63 EMBPRLSPDRAVL 78

## RESULT 10

H96700  
protein F12A21.5 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: H96700  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luross, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: H96700  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-464 <STO>  
A:Cross-references: GB:AE005173; NID:g11072032; PIDN:AAG28911.1; GSPDB:GN00141  
C:Genetics:  
A:Gene: F12A21.5  
A:Map position: 1

Query Match 44.9%; Score 44; DB 2; Length 464;

Best Local Similarity 64.3%; Pred. No. 37;  
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 2 KAPHLSPD--PQ 13  
I I I I I I I I  
Db 81 KRAHLNIPDLNPQ 94

## RESULT 11

C38128  
epithelin/granulin precursor - mouse  
N:Alternate names: acrogranin; PC-cell-derived growth factor  
C:Species: Mus musculus (house mouse)  
C>Date: 10-Jul-1992 #sequence\_revision 10-Jul-1992 #text\_change 20-Aug-1999  
C:Accession: C38128; S32503; I49468; A46705  
R:Plowman, G.D.; Green, J.M.; Neuberger, M.G.; Buckley, S.D.; McDonald, V.L.; Todaro, G.  
J. Biol. Chem. 267, 13073-13078, 1992  
A:Title: The epithelin precursor encodes two proteins with opposing activities on epithe  
A:Reference number: A38128; MUID:92317004; PMID:1618805  
A:Accession: C38128  
A:Molecule type: mRNA  
A:Residues: 1-589 <PLO>  
A:Cross-references: GB:X63231; NID:g50851; PIDN:CAAM4197.1; PID:g50852  
R:Baba, T.; Nemoto, H.; Watanabe, K.; Arai, Y.; Gerton, G.L.  
FEBS Lett. 322, 89-94, 1993  
A:Title: Exon/intron organization of the gene encoding the mouse epithelin/granulin prec  
A:Reference number: S32503; MUID:93245991; PMID:8482392  
A:Accession: S32503  
A:Molecule type: DNA  
A:Residues: 18-349, 'L', 351-589 <BAB>  
R:Baba, T.; Hoff, H.B.  
Mol. Reprod. Dev. 34, 233-243, 1993  
A:Title: Acrogranin, an acrosomal cysteine-rich glycoprotein, is the precursor of the gro  
A:Reference number: I48141; MUID:93228994; PMID:8471244  
A:Accession: I49468  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-250, 'L', 252-253, 'V', 255-349, 'L', 351-401, 'SA', 404-589 <RES>  
A:Cross-references: GB:M86736; NID:g191766; PID:AAA37191.1; PID:g191767  
R:Zhou, J.; Gao, G.; Crab, J.W.; Serrero, G.  
J. Biol. Chem. 268, 10863-10869, 1993

A:Title: Purification of an autocrine growth factor homologous with mouse epithelin  
A:Reference number: A46705; MUID:93266526; PMID:8496151  
A:Accession: A46705  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 18-19, 'X', 21-25, 'X', 27-29, 'XX', 32, 'XXX', 119-127, 152-154, 'DXK', 158-161,  
C:Superfamily: granulin

Query Match 44.9%; Score 44; DB 2; Length 589;

Best Local Similarity 55.6%; Pred. No. 48;  
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 EKAPHLSPDQALKRD 18  
I I I I I I I I  
Db 344 KVIAPRRLPDPQLKSD 361

## RESULT 12

G75563  
phosphate acetyltransferase (PC 2.3.1.8) - Deinococcus radiodurans (strain RI)  
C:Species: Deinococcus radiodurans  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jun-2000  
C:Accession: G75563  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, F.  
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans RI  
A:Reference number: A75250; MUID:20036896; PMID:10567266  
A:Accession: G75563  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-722 <WHI>  
A:Cross-references: GB:AE001870; GB:AE000513; NID:g6457732; PIDN:AAF09663.1; PID:g  
A:Experimental source: strain RI  
C:Genetics:  
A:Gene: DR0073  
A:Map position: 1  
C:Superfamily: phosphate acetyltransferase pta  
C:Keywords: acyltransferase; coenzyme A

Query Match 44.9%; Score 44; DB 2; Length 722;

Best Local Similarity 56.2%; Pred. No. 60;  
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 KAPHLSPDQALKR 17  
I I I I I I I I  
Db 75 RAVAHLTTPDPIPLTR 90

## RESULT 13

F82984  
hypothetical protein PA5299 [imported] - Pseudomonas aeruginosa (strain PAO1)  
C:Species: Pseudomonas aeruginosa  
C>Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: F82984  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.;  
Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic  
A:Reference number: A82950; MUID:20437337; PMID:10984043  
A:Accession: F82984  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-619 <STO>  
A:Cross-references: GB:AE004942; GB:AE004091; NID:g9951607; PIDN:AAG08684.1; GSPDB:  
A:Experimental source: strain PAO1  
C:Genetics:  
A:Gene: PA5299

Query Match 44.4%; Score 43.5; DB 2; Length 619;  
Best Local Similarity 45.5%; Pred. No. 61;

Search completed: July 8, 2003, 16:29:29  
Job time: 21.2436 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2003, 16:16:30 ; Search time 9.25641 Seconds  
(without alignments)  
85.136 Million cell updates/sec

Title: US-09-824-647-6

Perfect score: 98

Sequence: 1 EKAPAHLSLPPQALKRDV 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	98	100.0	593	1 GRN_HUMAN	P28799 h granulins
2	63	64.3	588	1 GRN_RAT	P23785 r granulins
3	50	51.0	589	1 GRN_MOUSE	P28798 mus musculus
4	43	43.9	990	1 NFRA_ECOLI	P31600 escherichia
5	42	42.9	129	1 SECG_PSEAE	Q9hrv52 pseudomonas
6	42	42.9	337	1 OPXK_MOUSE	O35214 mus musculus
7	42	42.9	462	1 WDR8_MOUSE	Q9jmg98 mus musculus
8	41	41.8	393	1 CYSK_YEAST	P53206 saccharomyc
9	41	41.8	397	1 CD4_ERIPA	Q08339 erythrocybu
10	41	41.8	410	1 ARR2_RAT	P29067 rattus norv
11	41	41.8	438	1 EIBL_ADECT	P14266 canine aden
12	41	41.8	610	1 BINA_HUMAN	O15213 homo sapien
13	41	41.8	710	1 ETF2_VACCA	O57220 vaccinia vi
14	41	41.8	710	1 ETF2_VACCC	P20635 vaccinia vi
15	41	41.8	710	1 ETF2_VACCV	Q9jfg93 vaccinia vi
16	41	41.8	710	1 ETF2_VACCV	P20636 vaccinia vi
17	41	41.8	710	1 ETF2_VARV	P33806 variola vir
18	41	41.8	745	1 ALD_HUMAN	P33897 homo sapien
19	41	41.8	910	1 SCL5_YEAST	P22224 saccharomyc
20	40.5	41.3	408	1 YOC2_YEAST	P25040 saccharomyc
21	40	40.8	203	1 CTF1_MOUSE	Q60753 mus musculus
22	40	40.8	203	1 CTF1_RAT	Q63086 rattus norv
23	40	40.8	282	1 3MG2_ECOLI	P04395 escherichia
24	40	40.8	323	1 CG96_HUMAN	Q9x3a4 homo sapien
25	40	40.8	350	1 RUVB_TREPA	P96115 treponema p
26	40	40.8	384	1 BOP1_HUMAN	Q9ubrl1 homo sapien
27	40	40.8	550	1 SYR_MYCTU	Q10609 mycobacteri
28	40	40.8	814	1 POL_IPMAI	P13894 mouse intra
29	40	40.8	850	1 NRG2_HUMAN	O14511 homo sapien
30	40	40.8	868	1 NRG2_RAT	O35569 rattus norv
31	40	40.8	3321	1 PCN2_HUMAN	O95613 homo sapien
32	39	39.8	326	1 FRA2_HUMAN	P13408 homo sapien
33	39	39.8	326	1 FRA2_MOUSE	P47930 mus musculus

34	39	39.8	327	1 FRA2_RAT	P51145 rattus norv
35	39	39.8	442	1 GLXD_RHIME	O87392 rhizobium m
36	39	39.8	442	1 TIG_CHLTR	O84713 chlamydia t
37	39	39.8	711	1 ETF2_MTXVL	Q9q8k4 myxoma viru
38	39	39.8	880	1 ARG2_YEAST	P05085 saccharomyc
39	39	39.8	1403	1 NID2_MOUSE	O88322 mus musculus
40	39	39.8	1748	1 POLR_ELV	P35928 erysimum la
41	39	39.8	2108	1 NOT1_YEAST	P12555 saccharomyc
42	39	39.8	2133	1 FAS_PIG	P12263 sus scrofa
43	38.5	39.3	355	1 AMBP_PLEPL	P36992 pleuronecte
44	38	38.8	268	1 ESL_HUMAN	P30042 homo sapien
45	38	38.8	296	1 TRA2_THIFE	P35884 thioabacillu

## ALIGNMENTS

## RESULT 1

GRN\_HUMAN  
ID GRN\_HUMAN STANDARD; PRT; 593 AA.  
AC P28799; P23781; P23782; P23783; P23784; Q9BWE7;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Granulins precursor (Acrogranin) [Contains: Paragranulin; Granulin 1  
(Granulin G); Granulin 2 (Granulin F); Granulin 3 (Granulin B);  
Granulin 4 (Granulin A); Granulin 5 (Granulin C); Granulin 6 (Granulin  
D); Granulin 7 (Granulin E)].  
GN GRN.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC TISSUE=Bone marrow;  
RX MEDLINE=92179253; PubMed=1542665;  
RA Bhandari V., Palfrey R.G.E., Bateman A.;  
RT "Isolation and sequence of the granulin precursor cDNA from human  
bone marrow reveals tandem cysteine-rich granulin domains.";  
Proc. Natl. Acad. Sci. U.S.A. 89:1715-1719(1992).  
RL [2]  
RN REVISIONS; SEQUENCE FROM N.A.  
RX MEDLINE=93038704; PubMed=1417868;  
RA Bhandari V., Bateman A.;  
RT "Structure and chromosomal location of the human granulin gene.";  
Biochem. Biophys. Res. Commun. 188:57-63(1992).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RX MEDLINE=92317004; PubMed=1618805;  
RA Plovan G.D., Green I.M., Neubauer M.G., Buckley S.D., McDonald V.L.,  
Todor G.J., Shoyab M.;  
RT "The epithelin precursor encodes two proteins with opposing activities  
on epithelial cell growth.";  
J. Biol. Chem. 267:13073-13078(1992).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX Yu W., Gibbs R.A.;  
RT Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
RC TISSUE=Cervix, and Lung;  
RX Strausberg R.;  
RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE OF 206-233; 281-336; 364-396 AND 442-447.  
RC TISSUE=Leukocyte;  
RX MEDLINE=91097544; PubMed=2268320;  
RA Bateman A., Belcourt D.R., Bennett H.P., Lazure C., Solomon S.;  
RT "Granulins, a novel class of peptide from leukocytes.";  
Biochem. Biophys. Res. Commun. 173:1161-1168(1990).  
RL





FT PEPTIDE 361 413 GRANULIN 5.  
 FT PEPTIDE 438 492 GRANULIN 6.  
 FT PEPTIDE 512 567 GRANULIN 7.  
 FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 372 372 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 525 525 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 201 201 S -> FP (IN REF. 2).  
 FT CONFLICT 307 308 TK -> SB (IN REF. 4).  
 FT CONFLICT 324 324 Q -> T (IN REF. 4).  
 FT CONFLICT 388 388 M -> I (IN REF. 2).  
 SQ SEQUENCE 588 AA; 63369 MW; 113D434F7E059B31 CRC64;

Query Match 64.38; Score 63; DB 1; Length 588;  
 Best Local Similarity 68.4; Pred. No. 0.021;  
 Matches 13; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 EKAPAHLSLPPDQALKRDV 19  
 DB 343 KKVTSLSLPPDQILKNDV 361

RESULT 3  
 GRN\_MOUSE STANDARD; PRT; 589 AA.  
 AC P28798;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Granulins precursor (Acrogranin) [Contains: Granulin 1; Granulin 2; Granulin 3; Granulin 4; Granulin 5; Granulin 6; Granulin 7].  
 GN GRN.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93245991; PubMed=8482392;  
 RA Baba T., Nemoto H., Watanabe K., Arai Y., Gerton G.L.;  
 RT "Exon/Intron organization of the gene encoding the mouse  
 epithelin/granulin precursor (acrogranin).";  
 RL FEBS Lett. 322:89-94(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92317004; PubMed=1618805;  
 RA Plowman G.D., Green I.M., Neubauer M.G., Buckley S., McDonald V.L.,  
 RA Todaro G.I., Shoyab M.;  
 RT "The epithelin precursor encodes two proteins with opposing  
 activities on epithelial cell growth.";  
 RL J. Biol. Chem. 267:13073-13078(1992).  
 CC -1- FUNCTION: GRANULINS HAVE POSSIBLE CYTOKINE-LIKE ACTIVITY. THEY MAY  
 PLAY A ROLE IN INFLAMMATION, WOUND REPAIR, AND TISSUE REMODELING.  
 CC -1- TISSUE SPECIFICITY: OBISQUITOUS.  
 CC -1- PFM: GRANULINS ARE DISULFIDE BRIDGED.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL; D16195; BAA03736.1; -  
 DR EMBL; M86736; AAA37191.1; -  
 DR EMBL; X62321; CAA44197.1; -  
 DR MGI; 95832; Grn.  
 DR InterPro; IPR000118; Granulin.  
 DR Pfam; PF00396; granulin; 7.  
 DR SMART; SM00277; GRN; 7.  
 DR PROSITE; PS00799; GRANULINS; 7.  
 KW Cytokine; Repeat; Glycoprotein; Signal.

FT SIGNAL 1 17  
 FT CHAIN 18 589  
 FT PEPTIDE 758 7113  
 FT PEPTIDE 7122 7178  
 FT PEPTIDE 205 260  
 FT PEPTIDE 280 334  
 FT PEPTIDE 362 7414  
 FT PEPTIDE 440 7493  
 FT PEPTIDE 7517 7568  
 FT CARBOHYD 38 38  
 FT CARBOHYD 263 263  
 FT CARBOHYD 373 373  
 FT CARBOHYD 526 526  
 FT CONFLICT 350 350  
 SQ SEQUENCE 589 AA; 63458 MW; 1DE8229C413CB787 CRC64;

Query Match 51.08; Score 50; DB 1; Length 589;  
 Best Local Similarity 61.18; Pred. No. 2.4;  
 Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 EKAPAHLSLPPDQALKRD 18  
 DB 344 KKVIAPLRLPPDQILKSD 361

RESULT 4  
 NFRA\_ECOLI STANDARD; PRT; 990 AA.  
 AC P31600;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Bacteriophage N4 adsorption protein A precursor.  
 GN NFRA OR B0568.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12;  
 RX MEDLINE=94042874; PubMed=8226649;  
 RA Kiino D.R., Singer M.S., Rothman-Denes L.B.;  
 RT "Two overlapping genes encoding membrane proteins required for  
 bacteriophage N4 adsorption.";  
 RL J. Bacteriol. 175:7081-7085(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:1453-1474(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Chung E., Allen E., Araujo R., Aparicio A., Davis K., Duncan M.,  
 RA Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O., Lew H., Lin D.,  
 RA Namath A., Oefner P., Roberts D., Schramm S., Davis R.W.;  
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12;  
 RX MEDLINE=97061202; PubMed=8905232;  
 RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,  
 RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kazimoto K.,  
 RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,  
 RA Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,  
 RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,  
 RA Yano M., Horuchi T.;  
 RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome

RT corresponding to the 12.7-28.0 min region on the linkage map.";  
RL DNA Res. 3:137-155(1996).  
CC -1- FUNCTION: REQUIRED FOR BACTERIOPHAGE N4 ADSORPTION. SERVES AS THE  
CC PHAGE RECEPTOR.  
CC  
CC -1- SUBCELLULAR LOCATION: Outer membrane.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; L16945; AAC36849.1; -  
DR EMBL; AE000161; NAC73669.1; -  
DR EMBL; U82598; AAB40766.1; -  
DR EMBL; D90699; BAA35202.1; -  
DR PIR; B49351; B49351.  
DR Ecogen; EG11740; nfrA.  
DR InterPro; IPR001440; TPR.  
DR Pfam; PF00515; TPR; 1.  
KW Phage recognition; Outer membrane; Signal; Complete proteome.  
FT SIGNAL 1 27 POTENTIAL.  
FT CHAIN 28 990 BACTERIOPHAGE N4 ADSORPTION PROTEIN A.  
SQ SEQUENCE 990 AA; 111307 MW; AAA616665D1C2E72 CRC64;  
  
Query Match 43.9%; Score 43; DB 1; Length 990;  
Best Local Similarity 56.2%; Pred. No. 54;  
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;  
  
QY 4 PAHLSPDPQALKRDV 19  
DB 671 PAHGLPDPALIRQL 686  
  
RESULT 5  
SECG\_PSEAE  
ID SECG\_PSEAE STANDARD; PRT; 129 AA.  
AC Q9HV52;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DE Protein-export membrane protein secg.  
GN SECG OR PA4747.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 15692 / PA01;  
RX MEDLINE=20437337; PubMed=10984043;  
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
RA Hickey M.J., Brinkman F.S.L., Hurnagale W.O., Kowalik D.J., Lagrou M.,  
RA Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,  
RA Brody L.L., Goltzy S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;  
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
RT opportunistic pathogen."  
RL Nature 406:959-964(2000).  
CC -1- FUNCTION: INVOLVED IN PROTEIN EXPORT. PARTICIPATES IN A EARLY  
CC EVENT OF PROTEIN TRANSLOCATION (BY SIMILARITY).  
CC  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
CC (By similarity).  
CC  
CC -1- SIMILARITY: BELONGS TO THE SECG FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; AE004888; AAG08133.1; -  
DR InterPro; IPR004692; SecG.  
DR TIGRFAMs; TIGR00810; secG; 1.  
KW Protein transport; Translocation; Transmembrane; Inner membrane;  
KW Complete proteome.  
FT TRANSMEM 5 25 POTENTIAL.  
FT TRANSMEM 56 75 POTENTIAL.  
SQ SEQUENCE 129 AA; 13207 MW; C1EF8BB2A3D52634 CRC64;  
  
Query Match 42.9%; Score 42; DB 1; Length 129;  
Best Local Similarity 47.4%; Pred. No. 8.3;  
Matches 9; Conservative 4; Mismatches 4; Indels 2; Gaps 1;  
  
QY 1 EKAPA--HLSPDPQALKR 17  
DB 77 EKSDALQHTGLPDPVLEQ 95  
  
RESULT 6  
OPSX\_MOUSE  
ID OPSX\_MOUSE STANDARD; PRT; 337 AA.  
AC Q35214;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Visual pigment-like receptor peropsin.  
GN RRR.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC MEDLINE=97420780; PubMed=9275222;  
RX Sun H., Gilbert D.J., Copeland N.G., Jenkins N.A., Nathans J.;  
RT "Peropsin, a novel visual pigment-like protein located in the apical  
RT microvilli of the retinal pigment epithelium."  
RL Proc. Natl. Acad. Sci. U.S.A. 94:9893-9898(1997).  
CC -1- FUNCTION: MAY PLAY A ROLE IN RPE PHYSIOLOGY EITHER BY DETECTING  
CC LIGHT DIRECTLY OR BY MONITORING THE CONCENTRATION OF RETINOIDS OR  
CC OTHER PHOTORECEPTOR-DERIVED COMPOUNDS.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
CC -1- TISSUE SPECIFICITY: FOUND ONLY IN THE EYE, WHERE IT IS LOCALIZED  
CC TO THE RETINAL PIGMENT EPITHELIUM (RPE). IN THE RPE IT IS  
CC LOCALIZED TO THE MICROVILLI THAT SURROUND THE PHOTORECEPTOR OUTER  
CC SEGMENTS.  
CC  
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
CC  
CC OPSIN SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; AF012271; AAC53344.1; -  
DR HSP; P02699; IBOJ.  
DR MG; MG1.1097709; Rrh.  
DR InterPro; IPR000276; GPCR\_Rhodpsn.  
DR InterPro; IPR001760; Opsin.  
DR Pfam; PF00001; 7tm\_1; 1.  
DR PROSITE; PS00237; G\_PROTEIN\_RECEPTOR\_F1\_1; 1.  
DR PROSITE; PS00262; G\_PROTEIN\_RECEPTOR\_F1\_2; 1.  
DR PROSITE; PS00238; OPSIN; 1.  
KW G-protein coupled receptor; Transmembrane; Glycoprotein.  
FT DOMAIN 1 26 EXTRACELLULAR.  
FT TRANSMEM 27 49 1 (POTENTIAL).

```
FT DOMAIN 50 61 CYTOPLASMIC.
FT TRANSSEM 62 87 2 (POTENTIAL).
FT DOMAIN 88 101 EXTRACELLULAR.
FT TRANSSEM 102 121 3 (POTENTIAL).
FT DOMAIN 122 140 CYTOPLASMIC.
FT TRANSSEM 141 164 4 (POTENTIAL).
FT DOMAIN 165 188 EXTRACELLULAR.
FT TRANSSEM 189 212 5 (POTENTIAL).
FT DOMAIN 213 240 CYTOPLASMIC.
FT TRANSSEM 241 264 6 (POTENTIAL).
FT DOMAIN 266 272 EXTRACELLULAR.
FT TRANSSEM 273 297 7 (POTENTIAL).
FT DOMAIN 298 337 CYTOPLASMIC.
FT CARBOHYD 9 9 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 182 182 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 98 175 POTENTIAL.
FT BINDING 284 284 RETINAL CHROMOPHORE.
SQ SEQUENCE 337 AA; 37208 MW; 353E4E2278EDBFA6 CRC64;

Query Match 42.9%; Score 42; DB 1; Length 337;
Best Local Similarity 38.9%; Pred. No. 24;
Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 KAPAHSLPDPQALKRDV 19
DB 310 KCQPHLAVPEFSLPMDM 327

RESULT 7
WDR8_MOUSE STANDARD; PRT; 462 AA.
AC Q9JMG9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE WD-repeat protein 8.
GN WDR8 OR D057.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21295046; PubMed=11401440;
RA Koshizuka Y., Ikegawa S., Sano M., Nakamura K., Nakamura Y.;
RT "Isolation, characterization, and mapping of the mouse and human WDR8
RT genes, members of a novel WD-repeat gene family.";
RL Genomics 72:252-259(2001).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- TISSUE SPECIFICITY: Ubiquitous.
CC -1- SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB034911; BAA92311.1;
CC MGD; MGI:1891749; Wdr8.
CC InterPro; IPR001680; WD40.
CC Pfam; PF00400; WD40; 3.
CC SMART; SM00320; WD40; 2.
CC PROSITE; PS00678; WD_REPEATS_1; FALSE NEG.
CC PROSITE; PS50082; WD_REPEATS_2; FALSE NEG.
CC PROSITE; PS50294; WD_REPEATS_REGION; FALSE NEG.
CC Repeat; WD repeat.
CC Repeat 46 86 WD 1.
CC REPEAT 89 129 WD 2.
CC REPEAT 176 210 WD 3.
CC REPEAT 221 260 WD 4.
FT
```

```
FT REPEAT 328 369 WD 5.
FT REPEAT 371 410 WD 6.
SQ SEQUENCE 462 AA; 52026 MW; 0142524BB1801D51 CRC64;

Query Match 42.9%; Score 42; DB 1; Length 462;
Best Local Similarity 47.4%; Pred. No. 34;
Matches 9; Conservative 3; Mismatches 3; Indels 4; Gaps 1;

QY 1 ERAP----AHLSPDPQAL 15
DB 279 EKSPLLGLGLHLSFPFPPAM 297

RESULT 8
CYSK_YEAST STANDARD; PRT; 393 AA.
AC P53206;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative cysteine synthase (EC 4.2.99.8) (O-acetylserine
DE sulfhydrylase) (O-acetylserine (Thiol)-lyase) (CSASE).
GN YGR012W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-S288C;
RX MEDLINE-97435481; PubMed=9290212;
RA Rieger M., Brueckner M., Schaefer M., Mueller-Auer S.;
RT "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae
RT chromosome VII.";
RL Yeast 13:1077-1090(1997).
CC -1- CATALYTIC ACTIVITY: O3-acetyl-L-serine + H(2)S = L-cysteine +
CC acetate.
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- PATHWAY: Cysteine biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE CYSTEINE SYNTHASE/CYSTATHIONINE BETA-
CC SYNTHASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z72797; CAA96995.1;
CC HSSP; P35520; 1JBQ.
CC SGP; S0003244; YGR012W.
CC InterPro; IPR001926; B6_enzyme_beta.
CC InterPro; IPR001216; Cys_synthase.
CC Pfam; PF00291; PALP; 1.
CC PROSITE; PS00901; CYS_SYNTHASE; 1.
CC Hypothetical protein; Lyase; Cysteine biosynthesis;
CC Pyridoxal phosphate.
CC BINDING 86 86
CC SEQUENCE 393 AA; 42801 MW; 76C73396D77B69AC CRC64;

Query Match 41.8%; Score 41; DB 1; Length 393;
Best Local Similarity 53.8%; Pred. No. 41;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 5 AHLSPDPQALKR 17
DB 134 AHLSLPDPTSLER 146

RESULT 9
CD4_ERYPA
```

ID CD4\_ERYPA STANDARD; PRT; 397 AA.

AC Q08339;

DT 01-FEB-1995 (Rel. 31, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE T-cell surface glycoprotein CD4 (T-cell surface antigen T4/Leu-3) (Fragment).

DE CD4

GN Erythrocybus patas (Red guenon) (Hussar).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Erythrocybus.

OC NCBI\_TaxID=9538;

OX [1]

RN SEQUENCE FROM N.A.

RP TISSUE=Blood;

RC MEDLINE=93049640; PubMed=1425921;

RX Fomsgaard A., Hirsch V.M., Johnson P.R.;

RA "Cloning and sequences of primate CD4 molecules: diversity of the cellular receptor for simian immunodeficiency virus/human immunodeficiency virus.";

RT Eur. J. Immunol. 22:2973-2981(1992).

RL

CC -!- FUNCTION: ACCESSORY PROTEIN FOR MHC CLASS-II ANTIGEN/T-CELL RECEPTOR INTERACTION. MAY REGULATE T-CELL ACTIVATION.

CC -!- SUBUNIT: ASSOCIATES WITH P56-LCK (BY SIMILARITY).

CC -!- SUBCELLULAR LOCATION: Type 1 membrane protein.

CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.

CC -!- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.

CC -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; X73324; CAA51750.1; -

DR HSP; P01730; IWIQ.

DR InterPro; IPR003006; Ig\_MHC.

DR InterPro; IPR003600; Ig\_Like.

DR InterPro; IPR003596; Ig\_v.

DR Pfam; PF00047; Ig\_2

DR SMART; SM00410; Ig\_Like; 1.

DR SMART; SM00406; Ig\_v; 1.

DR Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell; Immune response; Repeat; Lipoprotein; Palmitate.

KW

FT NON\_TER 1

FT DOMAIN <1 369 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 370 331 POTENTIAL.

FT DOMAIN 392 >397 CYTOPLASMIC (POTENTIAL).

FT DOMAIN <1 98 IG-LIKE V-TYPE DOMAIN.

FT DOMAIN 99 176 IG-LIKE C2-TYPE DOMAIN 1.

FT DOMAIN 177 290 IG-LIKE C2-TYPE DOMAIN 2.

FT DOMAIN 291 347 IG-LIKE C2-TYPE DOMAIN 3.

FT CARBOHYD 269 269 N-LINKED (GLCNAC. . .) (BY SIMILARITY).

FT CARBOHYD 298 298 N-LINKED (GLCNAC. . .) (BY SIMILARITY).

FT DISULFID 14 82 BY SIMILARITY.

FT DISULFID 128 157 BY SIMILARITY.

FT DISULFID 301 343 BY SIMILARITY.

FT LIPID 392 392 PALMITATE (BY SIMILARITY).

FT LIPID 395 395 PALMITATE (BY SIMILARITY).

FT NON\_TER 397 397

SQ SEQUENCE 397 AA; 44081 MW; 67887397A6B7EA4F CRC64;

Query Match 41.8%; Score 41; DB 1; Length 397;

Best Local Similarity 66.7%; Pred. No. 41;

Matches 10; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

OY 1 EKAPAHLSLPDPOAL 15

Db 249 EXLPLHLTL--PQAL 261

RESULT 10

ARR2\_RAT

ID ARR2\_RAT STANDARD; PRT; 410 AA.

AC P29067;

DT 01-DEC-1992 (Rel. 24, Created)

DT 01-DEC-1992 (Rel. 24, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Beta-arrestin 2 (Arrestin, beta 2).

DE ARRB2.

GN Rattus norvegicus (Rat).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OC NCBI\_TaxID=10116;

OX [1]

RN SEQUENCE FROM N.A.

RP STRAIN=Sprague-Dawley; TISSUE=Brain;

RC MEDLINE=92388146; PubMed=1517224;

RX Attramadal H., Arriza J.L., Aoki C., Dawson T.M., Codina J., Kwatra M.M., Snyder S.H., Caron M.G., Lefkowitz R.J.;

RA "Beta-arrestin2, a novel member of the arrestin/beta-arrestin gene family.";

RT J. Biol. Chem. 267:17882-17890(1992).

RL [2]

RN SEQUENCE OF 295-410 FROM N.A.

RP TISSUE=Pineal gland;

RC MEDLINE=94140898; PubMed=8308033;

RX Craft C.M., Whitmore D.H., Wiechmann A.F.;

RA "Cone arrestin identified by targeting expression of a functional family.";

RT J. Biol. Chem. 269:4613-4619(1994).

RL [3]

RN SEQUENCE OF 305-386 FROM N.A.

RP STRAIN=Sprague-Dawley; TISSUE=Retina;

RC MEDLINE=99145674; PubMed=9767391;

RX Komori N., Cain S.D., Roch J.-M., Miller K.E., Matsumoto H.;

RA "Differential expression of alternative splice variants of beta-arrestin-1 and -2 in rat central nervous system and peripheral tissues.";

RT Eur. J. Neurosci. 10:2607-2616(1998).

RL

CC -!- FUNCTION: REGULATES BETA-ADRENERGIC RECEPTOR FUNCTION. BETA-ARRESTINS SEEM TO BIND PHOSPHORYLATED BETA-ADRENERGIC RECEPTORS, THEREBY CAUSING A SIGNIFICANT IMPAIRMENT OF THEIR CAPACITY TO ACTIVATE G(S) PROTEINS.

CC -!- TISSUE SPECIFICITY: PREDOMINANTLY LOCALIZED IN NEURONAL TISSUES AND IN THE SPLEEN.

CC -!- SIMILARITY: BELONGS TO THE ARRESTIN FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; M91590; AAA74460.1; -

DR EMBL; U03627; AAA17551.1; -

DR EMBL; AF051457; AAC28617.1; -

DR HSP; P08168; 1CFL.

DR InterPro; IPR000698; Arrestin.

DR Pfam; PF00339; arrestin; 1.

DR Pfam; PF02752; arrestin\_C; 1.

DR PRINTS; PR00309; ARRESTIN.

DR PRODOM; PD002099; Arrestin; 1.

DR PROSITE; PS00295; ARRESTINS; 1.

KW Sensory transduction.

SQ SEQUENCE 410 AA; 46340 MW; 0DFA6A897C2B86BA CRC64;

Query Match 41.8%; Score 41; DB 1; Length 410;

Best Local Similarity 42.9%; Pred. No. 42;

Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

```

QY      6 HSLPDPQALKRDV 19
DB      353 HTLPRQSPAPREI 366

RESULT 11
EIBL_ADECT STANDARD; PRT; 438 AA.
AC P14266;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE E1B protein, large T-antigen (Early E1B 49 kDa protein).
OS Canine adenovirus type 2 (strain Toronto A 26-61).
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=69152;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE-90021176; PubMed-2800332;
RA Shibata R., Shinagawa M., Iida Y., Tsukiyama T.;
RL "Nucleotide sequence of E1 region of canine adenovirus type 2.";
Virology 172:460-467(1989).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z97184; CAB09994.1;
CC EMBL; BC000388; AAH00388.1;
CC Genew; HGNC:13923; C6orf11.
CC InterPro; IPR001680; WD40.
CC Pfam; PF00400; WD40; 1.
CC SMART; SM00320; WD40; 1.
CC PROSITE; PS00678; WD_REPEATS_1; FALSE_NEG.
CC PROSITE; PS00082; WD_REPEATS_2; 1.
CC PROSITE; PS0294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat. 234
FT REPEAT 193
FT REPEAT 235 272
FT REPEAT 274 312
FT REPEAT 315 354
FT REPEAT 357 396
FT REPEAT 399 436
SQ SEQUENCE 610 AA; 68041 MW; 0C1IDA8EDCBA3DEE CRC64;

Query Match 41.8%; Score 41; DB 1; Length 610;
Best Local Similarity 60.0%; Pred. No. 66;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY      1 EKAPAHLSLPDPQAL 15
DB      507 EKVPALICLDPRAL 521

RESULT 13
E1F2_VACCA STANDARD; PRT; 710 AA.
AC O57220;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Early transcription factor 82 kDa subunit (VETF large subunit).
GN MVAL18L.
OS Vaccinia virus (strain Ankara).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus
OX NCBI_TaxID=126794;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE-98263813; PubMed-9601507;
RA Antoine G., Scheiflinger F., Dörner F., Falkner F.G.;
RT "The complete genomic sequence of the modified vaccinia Ankara strain:
RT comparison with other orthopoxviruses.";
Virology 244:365-396(1998).
RN [2]
RN CHARACTERIZATION.
RX MEDLINE-90272720; PubMed-2190222;
RA Gershon P.D., Moss B.;
RT "Early transcription factor subunits are encoded by vaccinia virus
RT late genes.";
Proc. Natl. Acad. Sci. U.S.A. 87:4401-4405(1990).
CC -----
CC -1- FUNCTION: ACTS WITH RNA POLYMERASE TO INITIATE TRANSCRIPTION FROM
CC EARLY GENE PROMOTERS. A DNA-DEPENDENT ATPASE ACTIVITY IS
CC ASSOCIATED WITH VETF.
CC -1- SUBUNIT: HETERODIMER OF A 70 kDa AND A 82 kDa SUBUNIT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2003, 16:17:10 ; Search time 40.9231 Seconds  
(without alignments)  
95.665 Million cell updates/sec

Title: US-09-824-647-6  
Perfect score: 98  
Sequence: 1 EKAPAHLSLDPQALRKDV 19

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_21.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	98	100.0	413	Q9H8S1	Q9h8s1 homo sapien
2	50	51.0	589	Q9D2V3	Q9d2v3 mus musculus
3	49	50.0	700	Q9U0R0	Q9u0r0 homo sapien
4	47	48.0	158	Q52786	Q52786 acinetobact
5	47	48.0	324	Q9H9D1	Q9h9d1 homo sapien
6	47	48.0	828	Q9NVP9	Q9nvp9 homo sapien
7	47	48.0	915	Q8WX33	Q8wx33 homo sapien
8	46	46.9	174	Q8VEL6	Q8vel6 mus musculus
9	46	46.9	543	Q949X8	Q949x8 arabidopsis
10	46	46.9	1857	Q9LEK9	Q9lex9 arabidopsis
11	46	46.9	2143	Q9C8A6	Q9c8a6 arabidopsis
12	46	46.9	2151	Q9RVL2	Q9rvl2 arabidopsis
13	45	45.9	282	Q92P26	Q92p26 rhizobium m
14	45	45.9	299	Q8TKS0	Q8tkso methanosarc
15	45	45.9	374	Q55654	Q55654 synechocyst
16	44.5	45.4	116	Q93341	Q93341 struthio ca

17	44.5	45.4	560	13	Q08781	Q08781 gallus gall
18	44.5	45.4	918	13	Q9DG03	Q9dgq3 aegolius fu
19	44.5	45.4	918	13	Q9DGQ1	Q9dgq1 nymphicus h
20	44.5	45.4	918	13	Q9DGQ0	Q9dgq2 gallus gall
21	44.5	45.4	981	13	Q9DGQ2	Q9dgq2 gallus gall
22	44.5	45.4	1808	13	Q42142	Q42142 gallus gall
23	44	44.9	187	17	Q9YB33	Q9yb33 acropyrum p
24	44	44.9	196	16	Q8Y592	Q8y592 streptomyce
25	44	44.9	223	5	Q9XUX0	Q9xux0 caenorhabdi
26	44	44.9	248	17	Q9UXC2	Q9uxc2 sulfolobus
27	44	44.9	358	10	Q9LV10	Q9lv10 arabidopsis
28	44	44.9	367	10	Q9SPG3	Q9spg3 arabidopsis
29	44	44.9	464	10	Q9FXE4	Q9fxe4 arabidopsis
30	44	44.9	562	16	Q8XZ39	Q8xz39 ralstonia s
31	44	44.9	722	16	Q9RY77	Q9ry77 deinococcus
32	44	44.9	929	5	Q9N993	Q9n993 leishmania
33	43.5	44.4	619	16	Q9HTQ5	Q9htq5 pseudomonas
34	43.5	44.4	918	13	Q9DGQ4	Q9dgq4 aegolius fu
35	43	43.9	119	16	Q93J48	Q93j48 streptomyce
36	43	43.9	304	3	Q42806	Q42806 emericella
37	43	43.9	366	16	Q8YRJ1	Q8yrj1 anabaena sp
38	43	43.9	406	16	Q8U9E9	Q8u9e9 agrobacteri
39	43	43.9	449	4	Q9UEI6	Q9uei6 homo sapien
40	42.5	43.4	705	3	Q9Y8D5	Q9y8d5 cochliobolu
41	42	42.9	194	5	Q9NKR0	Q9nkr0 leishmania
42	42	42.9	223	5	Q27110	Q27110 trichomonas
43	42	42.9	227	4	Q96CM9	Q96cm9 homo sapien
44	42	42.9	425	5	Q19460	Q19460 caenorhabdi
45	42	42.9	430	16	Q91OU3	Q91ou3 streptomyce

ALIGNMENTS

RESULT 1

ID	Q9H8S1	PRELIMINARY;	PRT;	413 AA.
AC	Q9H8S1;			
DT	01-MAR-2001 (TREMBLrel. 16, Created)			
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	QVARC1001154 protein.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=OVARIAN CARCINOMA;			
RA	Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,			
RA	Nishikawa T., Nagai K., Sato H., Sugano S., Shiratori A., Sudo H.,			
RA	Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,			
RA	Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,			
RA	Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,			
RA	Yamanoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,			
RA	Ninomiya K., Iwayanagi T.,			
RT	*NEDO human cDNA sequencing project.*;			
RL	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AK023348; BAB14535.1; -			
DR	InterPro; IPR000118; Granulin.			
DR	InterPro; IPR001211; PhospholipaseA2.			
DR	Pfam; PF003396; granulin; 4.			
DR	SMART; SM00277; GRAN; 4.			
DR	PROSITE; PS00799; GRANULINS; UNKNOWN_4.			
DR	PROSITE; PS00118; PA2_HIS; UNKNOWN_1.			
DR	PROSITE; PS00118; PA2_HIS; UNKNOWN_1.			
DR	SEQUENCE 413 AA; 44132 MW; 0E3767A44BE314EC CRC64;			

Query Match 100.0%; Score 98; DB 4; Length 413;  
Best Local Similarity 100.0%; Pred. No. 7.9e-08;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKAPAHLSLDPQALRKDV 19  
|||||



Db 166 EKAPAHLSLPDPQALKRDV 184

RESULT 2

Q9D2V3 ID Q9D2V3 PRELIMINARY; PRT; 589 AA.

AC Q9D2V3

DT 01-JUN-2001 (TRENBLrel. 17, Created)

DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)

DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)

DE Adult male kidney cDNA, RIKEN full-length enriched library,  
clone:0610012H06, full insert sequence.

GN Mus musculus (Mouse).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=KIDNEY;

RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Aizawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
RA Hayashizaki Y.

RT "Functional annotation of a full-length mouse cDNA collection."

RL Nature 409:685-690(2001).

DR EMBL; AK018744; BAB31384.1; -.

DR MGI; MGI:95832; Grn.

DR InterPro; IPR000118; Granulin.

DR InterPro; IPR001211; PhospholipaseA2.

DR Pfam; PF00396; granulin; 7.

DR SMART; SM00277; GRAN; 7.

DR PROSITE; PS00799; GRANULINS; 7.

DR PROSITE; PS00118; PA2\_HIS; UNKNOWN\_1.

SQ SEQUENCE 589 AA; 63405 MW; 1D8229C413CA292 CRC64;

Query Match 51.0%; Score 50; DB 11; Length 589;  
Best Local Similarity 61.1%; Pred. No. 8.7;  
Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 EKAPAHLSLPDPQALKRDV 18

Db 344 KVIAPLRPLDPQLKSD 361

RESULT 3

Q9UQRO ID Q9UQRO PRELIMINARY; PRT; 700 AA.

AC Q9UQRO Q9UGCS;

DT 01-MAY-2000 (TRENBLrel. 13, Created)

DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)

DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)

DE SCML2 protein (DJ1129A6.1) (sex comb on midleg (Drosophila)-like  
2).

GN SCML2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=99265973; PubMed=10331946;

RA Montini E., Buchner G., Spalluto C., Andolfi G., Caruso A.,  
RA De Dumenil J.T., Trump D., Rocchi M., Ballabio A., Franco B.,  
RT "Identification of SCML2, a Second Human Gene Homologous to the  
RT Drosophila Sex comb on midleg (Scm): a New Gene Cluster on Xp22.2";  
RL Genomics 58:65-72(1999).

RN [2]

RP SEQUENCE OF 245-700 FROM N.A.

RA Heath P.;

RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; Y18004; CAB38943.1; -.

DR EMBL; AL096763; CAB65926.1; -.

DR InterPro; IPR004092; Mbt.

DR InterPro; IPR001660; SAM.

DR Pfam; PF02820; mbt; 2.

DR Pfam; PF00536; SAM; 1.

DR SMART; SM00454; SAM; 1.

SQ SEQUENCE 700 AA; 77256 MW; 05E086D4928DE73 CRC64;

Query Match 50.0%; Score 49; DB 4; Length 700;  
Best Local Similarity 43.8%; Pred. No. 15;  
Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 2 KAPAHLSLPDPQALKR 17

Db 607 EAPSYIAVDPSPVLKQ 622

RESULT 4

Q52786 ID Q52786 PRELIMINARY; PRT; 158 AA.

AC Q52786;

DT 01-JUN-1998 (TRENBLrel. 06, Created)

DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)

DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)

DE Hypothetical 18.4 kDa protein (Fragment).

OS Acinetobacter baumannii.

OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;

OC Acinetobacter.

OX NCBI\_TaxID=29430;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 17906;

RX MEDLINE=98321635; PubMed=9660198;

RA Bode C.E., Brabetz W., Brade H.;

RT "Cloning and characterization of 3-deoxy-D-manno-oct-2-ulosonic acid  
RT (Kdo) transferase genes (kdtA) from Acinetobacter baumannii and  
RT Acinetobacter baumannii."

RL Eur. J. Biochem. 254:404-412(1998).

DR EMBL; Z96927; CAB09655.1; -.

DR InterPro; IPR002106; AATRNA\_LigaseII.

DR PROSITE; PS00339; AA\_TRNA\_LIGASE\_II\_2; UNKNOWN\_1.

KW Hypothetical protein.

FT NON\_TER 158

SQ SEQUENCE 158 AA; 18399 MW; 81DA3A64E5152268 CRC64;

Query Match 48.0%; Score 47; DB 2; Length 158;  
Best Local Similarity 52.4%; Pred. No. 6.3;  
Matches 11; Conservative 3; Mismatches 5; Indels 2; Gaps 1;

QY 1 EKAPAHLSLPDPQALKRDV 19

Db 74 EKPLVHLILALPRPKVLRNV 94

RESULT 5

Q9H9D1 ID Q9H9D1 PRELIMINARY; PRT; 324 AA.

AC Q9H9D1

DT 01-MAR-2001 (TRENBLrel. 16, Created)

DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
 DE CDNA FLJ12840 fis, clone NT2RP2003277, weakly similar to NAM7  
 DE protein.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Isoqai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,  
 RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,  
 RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,  
 RA Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;  
 RT "NEDO human cDNA sequencing project."  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK022902; BAB14299.1; --  
 SQ SEQUENCE 324 AA; 35932 MW; EDDEA2D23EC145F4 CRC64;  
  
 Query Match 48.0%; Score 47; DB 4; Length 324;  
 Best Local Similarity 52.9%; Pred. No. 14;  
 Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;  
  
 QY 1 EKAPAHLSLPDPQALKR 17  
 I: | | | | | | | | | |  
 Db 175 ERPPVHDQLQDPRLKR 191  
  
 RESULT 6  
 Q9NVP9 PRELIMINARY; PRT; 828 AA.  
 ID Q9NVP9;  
 AC Q8WV66;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE CDNA FLJ10594 fis, clone NT2RP2004689.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Isoqai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,  
 RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,  
 RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,  
 RA Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;  
 RT "NEDO human cDNA sequencing project."  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK001456; BAA91701.1; --  
 SQ SEQUENCE 828 AA; 93316 MW; CAC902297F0EA129 CRC64;  
  
 Query Match 48.0%; Score 47; DB 4; Length 828;  
 Best Local Similarity 52.9%; Pred. No. 39;  
 Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;  
  
 QY 1 EKAPAHLSLPDPQALKR 17  
 I: | | | | | | | | | |  
 Db 679 ERPPVHDQLQDPRLKR 695  
  
 RESULT 7  
 Q8WX33 PRELIMINARY; PRT; 915 AA.  
 ID Q8WX33;  
 AC Q8WX33; 075120;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE BA479K20.2 (KIAA0625 protein) (Fragment).  
 GN BA479K20.2 OR KIAA0625.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Skuce C.;  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 125-915 FROM N.A.  
 RC TISSUE=BRAIN;  
 RX MEDLINE=98403980; PubMed=9734811;  
 RA Ishikawa K., Nagase T., Suyama M., Miyajima N., Tanaka A., Kotani H.,  
 RA Nomura N., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. X.  
 RT The complete sequences of 100 new cDNA clones from brain which can  
 RT code for large proteins in vitro."  
 RL DNA Res. 5:169-176(1998).  
 DR EMBL; AL159997; CAD13341.1; --  
 DR EMBL; AB014525; BAA31600.1; --  
 FT NON\_TER 1  
 SQ SEQUENCE 915 AA; 104041 MW; 3CAC94C174C9A677 CRC64;  
  
 Query Match 48.0%; Score 47; DB 4; Length 915;  
 Best Local Similarity 52.9%; Pred. No. 44;  
 Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;  
  
 QY 1 EKAPAHLSLPDPQALKR 17  
 I: | | | | | | | | | |  
 Db 766 ERPPVHDQLQDPRLKR 782  
  
 RESULT 8  
 Q8VEL6 PRELIMINARY; PRT; 174 AA.  
 ID Q8VEL6;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Hypothetical 20.0 kDa protein (Fragment).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC018234; AAH18234.1; --  
 KW Hypothetical protein.  
 FT NON\_TER 1  
 SQ SEQUENCE 174 AA; 20018 MW; 75250AD38F86A088 CRC64;  
  
 Query Match 46.9%; Score 46; DB 11; Length 174;  
 Best Local Similarity 50.0%; Pred. No. 10;  
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
  
 QY 4 PAHLSLPDPQALKRDV 19  
 I: | | | | | | | | | |  
 Db 158 PLHKLPDQTLQKEY 173  
  
 RESULT 9  
 Q949X8 PRELIMINARY; PRT; 543 AA.  
 ID Q949X8;  
 AC Q949X8;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Putative n-calpain-1 large subunit (Fragment).  
 GN F7A10.23.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

```

OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Pham P.K., Banh J., Banno F., Dale J.M.,
RA Goldsmith A.D., Jiang P.X., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Yamamura Y., Yu G., Yu S., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Kossema E., Lam B.,
RA Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Tracy S.E.,
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT *Full Length cDNA of gene F7A10.23 (GI:12323169).;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY050822; AAK92757.1;
DR InterPro: IPR002110; ANK.
DR SMART; SM00248; ANK; 1.
KW ANK repeat; Repeat.
FT NON_TER 543
SQ SEQUENCE 543 AA; 59467 MW; FB988FE03FD3AD48 CRC64;

Query Match 46.9%; Score 46; DB 10; Length 543;
Best Local Similarity 61.5%; Pred. No. 36;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 HSLSPDPQALKRD 18
DB 337 HSLTDPDSARRD 349

RESULT 10
Q9LEX9
ID Q9LEX9 PRELIMINARY; PRT; 1857 AA.
AC Q9LEX9
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Hypothetical 203.1 kDa protein.
GN T27115.10.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Choisne N., Robert C., Brottier P., Wincker P., Cattolico L.,
RA Artiguenave F., Saurin W., Weissenbach J., Mewes H.W., Rudd S.,
RA Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL358732; CAB94128.1;
DR InterPro: IPR000409; Beige_BEACH.
DR Pfam: PF02138; Beach; 1.
DR PRODOM: PD007848; Beige_BEACH; 1.
DR PROSITE: PSS0197; BEACH; 1.
KW Hypothetical protein.
SQ SEQUENCE 1857 AA; 203121 MW; D0E91EFCDF5B6F73 CRC64;

Query Match 46.9%; Score 46; DB 10; Length 1857;
Best Local Similarity 52.9%; Pred. No. 1.4e+02;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 KAPAHLSLDPQALKRD 18
DB 1336 EAPAHAPPAPPPPLRRD 1352

RESULT 11
Q9C8A6
ID Q9C8A6 PRELIMINARY; PRT; 2143 AA.

Q9C8A6;
01-JUN-2001 (TRENBLrel. 17, Created)
01-JUN-2001 (TRENBLrel. 17, Last sequence update)
01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE N-calpain-1 large subunit, putative.
F7A10.23.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Eguip, P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Militischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzer S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT *Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana.;
RL Nature 408:816-820(2000).
DR EMBL: AC027034; AAG51565.1;
DR InterPro: IPR001300; Protease_C2.
DR InterPro: IPR000169; SHprot_acsite.
DR Pfam: PF01067; Calpain_III; 1.
DR Pfam: PF00648; peptidase_C2; 1.
DR PRINTS; PR00704; CALPAIN.
DR SMART; SM00230; CysPC; 1.
DR PROSITE; PS00139; THIOLEPROTEASE_CYS; 1.
SQ SEQUENCE 2143 AA; 236779 MW; B91A15081AF26EC2 CRC64;

Query Match 46.9%; Score 46; DB 10; Length 2143;
Best Local Similarity 61.5%; Pred. No. 1.6e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 HSLSPDPQALKRD 18
DB 372 HSLTDPDSARRD 384

RESULT 12
Q8RVL2
ID Q8RVL2 PRELIMINARY; PRT; 2151 AA.
AC Q8RVL2
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Calpain-like protein.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=DEVELOPING SEED;
RX PubMed=11929961;
RA Lid S.E., Gruis D., Jung R., Lorentzen J.A., Ananiev E.,
RA Chamberlin M., Niu X., Meeley R., Nichols S., Olsen O.A.;
RT *The defective kernel 1 (dek1) gene required for aleurone cell

```

```

RT development in the endosperm of maize grains encodes a membrane
RT protein of the calpain gene superfamily.;
RL Proc. Natl. Acad. Sci. U.S.A. 99:5460-5465(2002).
DR EMBL; AY061803; AAL38186.1;
SQ SEQUENCE 2151 AA; 238262 MW; 50D571D7446A5609 CRC64;

Query Match 46.9%; Score 46; DB 10; Length 2151;
Best Local Similarity 61.5%; Pred. No. 1.6e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 HLSLPPQALKRD 18
   |||: |||: |||
Db 337 HLSITDPSAARD 349

RESULT 13
Q92P26 PRELIMINARY; PRT; 282 AA.
AC Q92P26;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE Putative transcription regulator protein.
GN R01970 OR SMC04315.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021.
RX MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kourty M., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
DR EMBL; AL591789; CAC46549.1; -.
DR InterPro; IPR000847; HTH_LysR.
DR InterPro; IPR005119; LysR_subst.
DR Pfam; PF00126; HTH_1; 1.
DR Pfam; PF03466; LysR_substrate; 1.
DR PROSITE; PS00044; HTH_LYSR_FAMILY; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 282 AA; 30379 MW; B9F223D63CE788B7 CRC64;

Query Match 45.9%; Score 45; DB 16; Length 282;
Best Local Similarity 80.0%; Pred. No. 25;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EKAPHLSLP 10
   |::|||::||
Db 241 ERAPHLGLP 250

RESULT 14
Q8TKS0 PRELIMINARY; PRT; 299 AA.
ID Q8TKS0;
AC Q8TKS0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical protein MA3328.
GN MA3328.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA Linton N., Naylor J., Stange-Thomann N., DeArelano K., Johnson R.,
RA Allen L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity."
RL Genome Res. 12:532-542(2002).
DR EMBL; AE011037; AAM06697.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 299 AA; 32345 MW; 68006B7448C5F426 CRC64;

Query Match 45.9%; Score 45; DB 17; Length 299;
Best Local Similarity 56.2%; Pred. No. 27;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 EKAPHLSLPQALK 16
   ||::||::|||
Db 281 EKYPALFEVPEQDLK 296

RESULT 15
ID Q55654 PRELIMINARY; PRT; 374 AA.
AC Q55654;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein slr0358.
GN SLR0358.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PCC 6803;
RA Tabata S.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96127529; PubMed=8590279;
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugita M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 64% to 92% of the genome."
RL DNA Res. 2:153-166(1995).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugita M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions."
RL DNA Res. 3:109-136(1996).
DR EMBL; D64000; BAA10174.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 374 AA; 42416 MW; 99D408A4D0BDA50F CRC64;

Query Match 45.9%; Score 45; DB 16; Length 374;
Best Local Similarity 50.0%; Pred. No. 35;

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2003, 16:08:55 ; Search time 35 seconds

(without alignments)  
53.300 Million cell updates/sec

Title: US-09-824-647-7

Perfect score: 74

Sequence: 1 ARRGTKCLREAPR 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*  
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	74	100.0	14	AAW85481	Human GP88 autocri
2	74	100.0	593	AAW8673	Granulin sequence.
3	74	100.0	593	AAW85475	Human GP88 autocri
4	74	100.0	593	AAE20521	Human granulin/epi
5	68	91.9	413	AAW94550	Human protein sequ
6	68	91.9	593	AAW14326	Human epithelin pr
7	68	91.9	621	AAW43971	Human cancer assoc
8	53	71.6	14	AAW85479	Mouse GP88 autocri
9	53	71.6	589	AAW14327	Mouse epithelin pr
10	53	71.6	589	AAW85474	Mouse GP88 autocri

11	53	71.6	589	23	AAE20520	Mouse granulin/epi
12	51	68.9	589	12	AAW14325	Rat epithelin prec
13	46	62.2	101	22	AAW48409	Propionibacterium
14	43	58.1	56	14	AAW41885	Granulin E. Homo
15	41	55.4	113	20	AAW29199	Amino acid sequenc
16	41	55.4	1083	23	ABW92875	Herbicidally activ
17	40.5	54.7	88	21	AAW40966	Zea mays protein f
18	40	54.1	55	22	AAW63381	Propionibacterium
19	40	54.1	64	22	AAW43048	Propionibacterium
20	40	54.1	76	22	AAW47116	Propionibacterium
21	40	54.1	255	22	ABW23868	Novel human diagno
22	40	54.1	422	19	AAW56275	Human nucleotide p
23	40	54.1	1156	20	AAW23754	Flavobacterium ker
24	39	52.7	83	22	AAW49674	Propionibacterium
25	39	52.7	255	20	AAW88397	Mouse neuro-growth
26	39	52.7	255	23	ABW72368	Murine protein iso
27	39	52.7	275	21	AAW52141	Mouse TANGO 125 (T
28	39	52.7	275	23	AAE21081	Mouse T125 (TANGO-
29	39	52.7	278	20	AAW88392	Mouse neuro-growth
30	39	52.7	278	22	AAE05356	Mouse Notch4-like
31	38.5	52.0	99	22	AAW49368	Propionibacterium
32	38	51.4	70	22	AAW48478	Propionibacterium
33	38	51.4	119	22	AAW23432	Novel human enzyme
34	38	51.4	119	22	AAW21798	Novel human neopla
35	38	51.4	131	19	AAW72069	HSV-2 strain SB5 C
36	38	51.4	137	22	AAW21670	Novel human neopla
37	38	51.4	182	22	AAW46132	Propionibacterium
38	38	51.4	302	19	AAW72214	HSV-2 strain SB5 C
39	38	51.4	306	20	AAW67755	HSV-2 VP22 protein
40	38	51.4	621	22	ABW30299	Novel human diagno
41	38	51.4	655	22	ABW24248	Novel human diagno
42	38	51.4	737	22	AAW39448	Human polypeptide
43	38	51.4	761	22	AAW39447	Human polypeptide
44	38	51.4	761	23	ABW84252	Human macro protei
45	38	51.4	770	22	ABW05609	Novel human diagno

#### ALIGNMENTS

RESULT 1  
AAW85481  
ID AAW85481 standard; Peptide; 14 AA.  
XX  
AC AAW85481;  
DT  
DT 15-MAR-1999 (first entry)  
XX  
DE Human GP88 autocrine growth factor antigenic peptide A14R.  
XX  
GP88; granulin; epithelin; human; growth factor; autocrine; tumour;  
KW cancer; viral infection; antagonist; therapy; diagnosis; antigen;  
KW antibody.  
XX  
OS Homo sapiens.  
XX  
PN WO9852607-A1.  
XX  
XX 26-NOV-1998.  
XX  
XX 22-MAY-1998; 98WO-US10555.  
XX  
PR 16-DEC-1997; 97US-0991862.  
PR 23-MAY-1997; 97US-0863079.  
XX  
(SERR/) SERRERO G.  
XX  
PI Serrero G;  
XX  
DR WPI; 1999-045276/04.  
XX  
PT Composition containing antagonist of growth factor GP88 - useful for  
treating cancer and viral diseases and also for diagnosing disease

PT from altered GP88 expression  
 XX Example 8; Page 45; 86pp; English.  
 PS This is the amino acid sequence of peptide A14R, comprising amino  
 CC acid residues A566-R579 of human GP88 (see AAW85475). GP88 is an 88  
 CC kDa glycoprotein autocrine growth factor that is expressed in a  
 CC tightly regulated manner in normal cells, is overexpressed and  
 CC unregulated in highly tumorigenic cells derived from normal cells,  
 CC and which acts as a stringently required growth stimulator for the  
 CC tumorigenic cells. A14R was used in an attempt to raise  
 CC neutralising antibodies to GP88 (see also AAW85480). Antagonists to  
 CC GP88, such as anti-GP88 antibodies, are used to treat diseases  
 CC associated with increased expression of GP88, particularly cancer  
 CC but also viral infections. Anti-GP88 antibodies can also be used  
 CC as diagnostic reagents and to deliver toxins or other compounds to  
 CC GP88-expressing cells.  
 XX  
 XX SQ Sequence 14 AA;  
 Query Match 100.0%; Score 74; DB 20; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 3.5e-06;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 ARRGTKCLRREAPR 14  
 DB 1 ARRGTKCLRREAPR 14  
 RESULT 2  
 AAR48673  
 ID AAR48673 standard; Protein; 593 AA.  
 XX  
 AC AAR48673;  
 XX  
 DT 22-APR-1994 (first entry)  
 XX  
 DE Granulin sequence.  
 XX  
 KW Granulin; keratinocytes; wound healing; inhibition; peptide;  
 KW granulocytes; leucocytes.  
 XX  
 OS Homo sapiens.  
 XX  
 PH Key Location/Qualifiers  
 FT Misc-difference 452  
 FT /note- "Valine encoded by ATG."  
 FT Misc-difference 539  
 FT /note- "Glycine encoded by CAG."  
 XX  
 PN WO9315195-A.  
 XX  
 PD 05-AUG-1993.  
 XX  
 PF 28-FEB-1992; 92WO-CA00089.  
 XX  
 PR 03-FEB-1992; 92US-0829233.  
 XX  
 PA (SOLO/) SOLOMON S.  
 XX  
 PI Solomon S;  
 DR WPI: 1993-320328/40.  
 DR N-PSDB; AAQ49052.  
 XX  
 PT New cystine rich granulin peptide(s) from leucocyte(s) - are  
 PT keratinocyte inhibitors useful topically for wound healing  
 PS Disclosure; Figure 4c; 53pp; English.  
 XX  
 CC The granulin inhibits keratinocytes and is useful in formulations  
 CC for promoting the healing of wounds.  
 XX

SQ Sequence 593 AA;  
 Query Match 100.0%; Score 74; DB 14; Length 593;  
 Best Local Similarity 100.0%; Pred. No. 0.00013;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 ARRGTKCLRREAPR 14  
 DB 566 ARRGTKCLRREAPR 579  
 RESULT 3  
 AAW85475  
 ID AAW85475 standard; Protein; 593 AA.  
 XX  
 AC AAW85475;  
 XX  
 DT 15-MAR-1999 (first entry)  
 XX  
 DE Human GP88 autocrine growth factor.  
 XX  
 KW GP88; granulin; epithillin; human; growth factor; autocrine; tumour;  
 KW cancer; viral infection; antagonist; therapy; diagnosis.  
 XX  
 OS Homo sapiens.  
 XX  
 PH Key Location/Qualifiers  
 FT Peptide 340..364  
 FT /note- "E19V peptide used to raise antibody"  
 FT Peptide 566..579  
 FT /note- "A14R peptide used to raise antibody"  
 XX  
 PN WO9852607-A1.  
 XX  
 PD 26-NOV-1998.  
 XX  
 PF 22-MAY-1998; 98WO-US10555.  
 XX  
 PR 16-DEC-1997; 97US-0991862.  
 PR 23-MAY-1997; 97US-0863079.  
 XX  
 PA (SERR/) SERRERO G.  
 XX  
 PI Serrero G;  
 XX  
 DR WPI: 1999-045276/04.  
 DR N-PSDB; AAW82825.  
 XX  
 PT Composition containing antagonist of growth factor GP88 - useful for  
 PT treating cancer and viral diseases and also for diagnosing disease  
 PT from altered GP88 expression  
 XX  
 PS Example 5; Fig 9A; 86pp; English.  
 XX  
 CC This is the amino acid sequence of human GP88, an 88 kDa  
 CC glycoprotein autocrine growth factor and epithillin/granulin  
 CC precursor that is expressed in a tightly regulated manner in normal  
 CC cells, is overexpressed and unregulated in highly tumorigenic cells  
 CC derived from normal cells, and which acts as a stringently required  
 CC growth stimulator for the tumorigenic cells. Inhibition of GP88  
 CC expression or action in the tumorigenic cells results in an  
 CC inhibition of the tumorigenic properties of the overproducing  
 CC cells. Antagonists to GP88 are used to treat diseases associated  
 CC with increased expression of GP88, particularly cancer but also  
 CC viral infections. Fragments of GP88 are used to raise specific  
 CC antibodies (used as antagonists, as diagnostic reagents and for  
 CC delivering toxins or other compounds to GP88-expressing cells) and  
 CC to screen for antibodies. Methods are provided for diagnosing  
 CC disease, or determining susceptibility to disease, resulting from  
 CC altered GP88 activity.  
 XX  
 XX SQ Sequence 593 AA;

Query Match 100.08; Score 74; DB 20; Length 593;  
 Best Local Similarity 100.08; Pred. No. 0.00013;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARRGTKCLRRREAPR 14  
 |||||  
 DB 566 ARRGTKCLRRREAPR 579

## RESULT 4

AAE20521  
 ID AAE20521 standard; Protein; 593 AA.

AC AAE20521;

DT 01-JUL-2002 (first entry)

DE Human granulin/epithelin precursor (GP88) protein.

XX Human; granulin precursor; GP88; cytostatic; tumorigenicity; tamoxifen;  
 KW antineoplastic; antioestrogen therapy; skin cancer.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 273

FT /note= "Encoded by AAG"

FT Region 346..364

FT /note= "Region used as antigen to develop anti-human  
 GP88 neutralising antibody"

XX US2002025543-A1.

XX 28-FEB-2002.

XX 15-JUN-2001; 2001US-0880842.

XX 23-MAY-1997; 97US-0863079.

XX 08-DEC-1999; 99US-0456886.

XX (SERR/) SERRERO G.

XX Serrero G;

XX WPI; 2002-267529/31.

XX N-PSDB; AAD32850.

XX Diagnosing tumorigenicity in a human, comprising obtaining a cell  
 sample, detecting GP88 in the cells, and determining the number of GP88  
 positive cells in the sample -

XX Disclosure; Fig 9B; 50pp; English.

XX The invention relates to a method for diagnosing tumorigenicity in a  
 human. The method comprises obtaining a biological sample containing  
 cells from the patient, detecting GP88 in the cells of the sample,  
 and determining the number of GP88 positive cells in the sample, and  
 determining the ratio of GP88 positive cells to the total number of cells  
 in the sample. The invention also relates to a method for determining  
 if a human patient is resistant to the antineoplastic effects of  
 antioestrogen therapy. The method is useful for diagnosing

XX tumorigenicity in a sample, such as blood, serum, plasma, urine, nipple  
 aspirate, cerebrospinal fluid, liver, kidney, breast, bone, brain, lung,  
 colon, or skin cancer. The method can be used to treat or prevent  
 re-occurrence of cancer in a patient, by administering tamoxifen if the  
 sample contains less than 10 % GP88, or less than 5 % GP88 positive  
 cells. The present sequence is human granulin/epithelin precursor (GP88)  
 protein.

XX Sequence 593 AA;

Query Match 100.08; Score 74; DB 23; Length 593;  
 Best Local Similarity 100.08; Pred. No. 0.00013;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARRGTKCLRRREAPR 14  
 |||||  
 DB 566 ARRGTKCLRRREAPR 579

## RESULT 5

AAAB94550  
 ID AAB94550 standard; Protein; 413 AA.

XX AC AAB94550;

XX 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:15310.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99TP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isoqai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602  
 full-length cDNAs defined in the specification, and for the detection  
 and/or diagnosis of the abnormality of the proteins encoded by the  
 full-length cDNAs -

XX Claim 8; SEQ ID 15310; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesising 5602  
 full-length cDNAs defined in the specification. Where a primer set  
 comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 to the complementary strand of a polynucleotide which comprises one of  
 the 5602 nucleotide sequences defined in the specification, where the  
 oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 of an oligonucleotide comprising a sequence complementary to the  
 complementary strand of a polynucleotide which comprises a 5'-end  
 sequence and an oligonucleotide comprising a sequence complementary to a  
 polynucleotide which comprises a 3'-end sequence, where the  
 oligonucleotide comprises at least 15 nucleotides and the combination of  
 the 5'-end sequence/3'-end sequence is selected from those defined in  
 the specification. The primer sets can be used in antisense therapy and  
 in gene therapy. The primers are useful for synthesising polynucleotides,  
 particularly full-length cDNAs. The primers are also useful for the  
 detection and/or diagnosis of the abnormality of the proteins encoded by  
 the full-length cDNAs. The primers allow obtaining of the full-length  
 cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
 AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 represent oligonucleotides, all of which are used in the exemplification  
 of the present invention.

XX Sequence 413 AA;

Query Match 91.9%; Score 68; DB 22; Length 413;



ET-1 is a bifunctional growth regulator, capable of stimulating the growth of some cell types while inhibiting the growth of others. ET-2 is functionally similar to ET-1 w.r.t. growth inhibitory

CC disorders, allergic reactions, graft versus host disease and organ  
 CC rejection, modulate haemostatic or thrombolytic activity, modulate  
 CC inflammation, cancers, cardiovascular disorders, neurological disease and  
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,  
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to  
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of  
 CC the present invention.

XX SQ Sequence 621 AA;

Query Match 91.9%; Score 68; DB 21; Length 621;

Best Local Similarity 92.9%; Pred. No. 0.0015;

Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AARGTKCLRREAPR 14  
 | |||||  
 DB 594 AARGTKCLRREAPR 607

# RESULT 8

AAW85479

ID AAW85479 standard; Peptide; 14 AA.

XX AC AAW85479;

XX DT 15-MAR-1999 (first entry)

XX DE Mouse GP88 autocrine growth factor antigenic peptide S14R.

XX GP88; granulins; epithelins; mouse; growth factor; autocrine; tumour;  
 KW cancer; viral infection; antagonist; therapy; diagnosis; antigen;  
 KW antibody.

XX OS Mus sp.

XX PN WO9852607-Al.

XX PD 26-NOV-1998.

XX PF 22-MAY-1998; 98WO-US10555.

XX PR 16-DEC-1997; 97US-0991862.

XX PR 23-MAY-1997; 97US-0863079.

XX PA (SERR/) SERRERO G.

XX PI Serrero G;

XX PN WPI; 1999-045276/04.

XX Composition containing antagonist of growth factor GP88 - useful for  
 PT treating cancer and viral diseases and also for diagnosing disease  
 PT from altered GP88 expression

XX Example 8; Page 45; 86pp; English.

XX This is the amino acid sequence of peptide S14R, comprising amino  
 CC acid residues S562-R575 of murine GP88 (see AAW85479). GP88 is an 88  
 CC kDa glycoprotein autocrine growth factor that is expressed in a  
 CC tightly regulated manner in normal cells, is overexpressed and  
 CC unregulated in highly tumorigenic cells derived from normal cells,  
 CC and which acts as a stringently required growth stimulator for the  
 CC tumorigenic cells. S14R was used to raise neutralising antibodies  
 CC to GP88. Antagonists to GP88, such as anti-GP88 antibodies, are  
 CC used to treat diseases associated with increased expression of  
 CC GP88, particularly cancer but also viral infections. Anti-GP88  
 CC antibodies can also be used as diagnostic reagents and to deliver  
 CC toxins or other compounds to GP88-expressing cells.

XX SQ Sequence 14 AA;

Query Match

Best Local Similarity 71.6%; Score 53; DB 20; Length 14;

Pred. No. 0.016;

Matches 9; Conservative. 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 RGTCKLRREAPR 14  
 |||||  
 DB 3 RGTCKLRKKIPR 14

# RESULT 9

AA14327

ID AA14327 standard; Protein; 589 AA.

XX AC AA14327;

XX DT 17-JAN-1992 (first entry)

XX DE Mouse epithelin precursor.

XX KW ET; growth regulation; inhibition; stimulation.

XX OS Mus musculus;

XX FH Key

XX FT Protein

XX FT Location/Qualifiers

XX FT /label= precursor

XX FT /note= "claim 21, page 55"

XX FT 280..335

XX FT /label= EP-1

XX FT /note= "claim 22, page 55"

XX FT 205..261

XX FT /label= EP-2

XX FT /note= "claim 23, page 55"

XX FT 59..114

XX FT /label= EP

XX FT /note= "claim 24, page 55"

XX FT 123..179

XX FT /label= EP

XX FT /note= "claim 25, page 55"

XX FT 362..416

XX FT /label= EP

XX FT /note= "claim 26, page 56"

XX FT 440..495

XX FT /label= EP

XX FT /note= "claim 27, page 56"

XX FT 515..570

XX FT /label= EP

XX FT /note= "claim 28, page 56"

XX PN WO9115510-A.

XX PD 17-OCT-1991.

XX PF 03-APR-1991; 91WO-US02321.

XX PR 13-MAR-1991; 91US-0083796.

XX PR 03-APR-1990; 90US-0504508.

XX (BRIM ) BRISTOL-MYERS SQUIB.

XX PI Shoyab M, Plowman GD;

XX WPI; 1991-325168/44.

XX N-PSDB; AAQ14340.

XX New cysteine-rich growth modulating proteins, epithelins - useful  
 PT as inhibitors of neoplastic cell growth and to promote wound  
 PT healing and treat psoriasis

XX Disclosure; Fig 23; 97pp; English.

XX ET-1 is a bifunctional growth regulator, capable of stimulating  
 CC the growth of some cell types while inhibiting the growth of others.  
 CC ET-2 is functionally similar to ET-1 w.r.t. growth inhibitory  
 CC bioactivity. In contrast, however, ET-2 is apparently not capable of

CC eliciting the growth stimulatory activity characteristic of ET-1 and,  
CC in fact, antagonises this ET-1 activity.  
CC See also AAQ14338-40, AAQ14952-53, AAR14328-9 and AAR15315-20.

XX Sequence 589 AA; DB 12; Length 589; Indels 0; Gaps 0;

Query Match 71.6%; Score 53; DB 12; Length 589; Indels 0; Gaps 0;

Best Local Similarity 75.0%; Pred. No. 0.6; Mismatches 2; Indels 1; Gaps 0;

Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RGTCLRRREAPR 14

Db 564 RGTCLRRKKIPR 575

AAW85474 standard; Protein; 589 AA.

AAW85474;

15-MAR-1999 (first entry)

Mouse GP88 autocrine growth factor.

GP88; granulatin; epithelin; mouse; growth factor; autocrine; tumour;

cancer; viral infection; antagonist; therapy; diagnosis.

Mus sp.

Key Location/Qualifiers

Misc-difference 8 /note= "encoded by CTG"

Misc-difference 54 /note= "encoded by AGC"

Misc-difference 377 /note= "encoded by TGA"

Peptide 208..219 /note= "P12T peptide used to raise antibody"

Peptide 344..362 /note= "K19T peptide, used to raise antibody"

Peptide 562..575 /note= "S14R peptide, used to raise antibody"

WO9852607-A1.

26-NOV-1998.

22-MAY-1998; 98WO-US10555.

16-DEC-1997; 97US-0991862.

23-MAY-1997; 97US-0863079.

(SERR/) SERRERO G.

Serrero G;

WPI; 1999-045276/04.

N-PSDB; AAV82824.

Composition containing antagonist of growth factor GP88 - useful for

treating cancer and viral diseases and also for diagnosing disease

from altered GP88 expression

Example 5; Fig 8A-D; 86pp; English.

This is the amino acid sequence of murine GP88, an 88 kDa

glycoprotein autocrine growth factor and epithelin/granulin

precursor that is expressed in a tightly regulated manner in normal

cells, is overexpressed and unregulated in highly tumorigenic cells

derived from normal cells, and which acts as a stringently required

growth stimulator for the tumorigenic cells. Inhibition of GP88

expression or action in the tumorigenic cells results in an

CC inhibition of the tumorigenic properties of the overproducing

cells. Murine GP88 cDNA (see AAV82824) was obtained from the highly

tumorigenic PC cell line. Antagonists to GP88 are used to treat

diseases associated with increased expression of GP88, particularly

cancer but also viral infections. Fragments of GP88 are used to

raise specific antibodies (used as antagonists, as diagnostic

reagents and for delivering toxins or other compounds to GP88-

expressing cells) and to screen for antibodies. Methods are

provided for diagnosing disease, or determining susceptibility to

disease, resulting from altered GP88 activity.

XX Sequence 589 AA;

Query Match 71.6%; Score 53; DB 20; Length 589;

Best Local Similarity 75.0%; Pred. No. 0.6;

Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RGTCLRRREAPR 14

Db 564 RGTCLRRKKIPR 575

AAE20520 standard; Protein; 589 AA.

AAE20520;

01-JUL-2002 (first entry)

Mouse granulin/epithelin precursor (GP88) protein.

Mouse; granulatin precursor; GP88; cytostatic; tumourigenicity; tamoxifen;

antineoplastic; antioestrogen therapy; skin cancer.

Mus sp.

Key Location/Qualifiers

Misc-difference 335 /label= "Unknown"

Region 344..362 /note= "Encoded by ATG"

Region /note= "Regions used as antigens to raise anti-GP88

antibodies"

Region 562..575 /note= "Regions used as antigens to raise anti-GP88

antibodies"

Misc-difference 586 /note= "Encoded by ACA"

US2002025543-A1.

28-FEB-2002.

15-JUN-2001; 2001US-0880842.

23-MAY-1997; 97US-0863079.

08-DEC-1999; 99US-0456886.

(SERR/) SERRERO G.

Serrero G;

WPI; 2002-267529/31.

N-PSDB; AAD32849.

Diagnosing tumorigenicity in a human, comprising obtaining a cell

sample, detecting GP88 in the cells, and determining the number of GP88

positive cells in the sample -

Disclosure; Fig 8; 50pp; English.

The invention relates to a method for diagnosing tumourigenicity in a

CC human. The method comprises obtaining a biological sample containing  
 CC cells from the patient, detecting GP88 in the cells of the sample, and  
 CC and determining the number of GP88 positive cells in the sample, and  
 CC determining the ratio of GP88 positive cells to the total number of cells  
 CC in the sample. The invention also relates to a method for determining  
 CC if a human patient is resistant to the antineoplastic effects of  
 CC antioestrogen therapy. The method is useful for diagnosing  
 CC tumourigenicity in a sample, such as blood, serum, plasma, urine, nipple  
 CC aspirate, cerebrospinal fluid, liver, kidney, breast, bone, brain, lung,  
 CC colon, or skin cancer. The method can be used to treat or prevent  
 CC re-occurrence of cancer in a patient, by administering tamoxifen if the  
 CC sample contains less than 10 % GP88, or less than 5 % GP88 positive  
 CC cells. The present sequence is mouse granululin/epithelin precursor (GP88)  
 CC protein.

XX SQ Sequence 589 AA;

Query Match 71.6%; Score 53; DB 23; Length 589;  
 Best Local Similarity 75.0%; Pred. No. 0.6;  
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 3 RGTCLRRREAPR 14  
 :|||||::||  
 Db 564 RGTCLRRKTKPR 575

RESULT 12

AAU14325  
 ID AAR14325 standard; Protein; 589 AA.

XX AC AAR14325;

XX DT 17-JAN-1992 (first entry)

XX DE Rat epithelin precursor.

XX EW ET; growth regulation; inhibition; stimulation.

XX OS Rattus rattus.

XX FH Key Location/Qualifiers

XX FT Protein 1..589

XX FT /label= precursor

XX FT /note= "claim 11, page 54"

XX FT Protein 280..335

XX FT /label= EP-1

XX FT /note= "claim 12, page 54"

XX FT Protein 205..261

XX FT /label= EP-2

XX FT /note= "claim 13, page 54"

XX FT Peptide 59..114

XX FT /label= EP

XX FT /note= "claim 14, page 54"

XX FT Peptide 123..179

XX FT /label= EP

XX FT /note= "claim 15, page 54"

XX FT Peptide 362..416

XX FT /label= EP

XX FT /note= "claim 16, page 54"

XX FT Peptide 440..495

XX FT /label= EP

XX FT /note= "claim 17, page 54"

XX FT Peptide 515..570

XX FT /label= EP

XX FT /note= "claim 18, page 55"

XX PN W09115510-A.

XX PD 17-OCT-1991.

XX PF 03-APR-1991; 91WO-US02321.

XX PR 13-MAR-1991; 91US-0083796.

PR 03-APR-1990; 90US-0504508.

XX PA (BRIM ) BRISTOL-MYERS SQUIB.

XX PI Shoyab M, Plozman GD;

XX DR WPI; 1991-325168/44.

XX DR N-PSDB; AAR14338.

XX New cysteine-rich growth modulating proteins, epithelins - useful  
 as inhibitors of neoplastic cell growth and to promote wound  
 healing and treat psoriasis

PS Disclosure; Fig 18; 97pp; English.

XX ET-1 is a bifunctional growth regulator, capable of stimulating  
 the growth of some cell types while inhibiting the growth of others.  
 CC ET-2 is functionally similar to ET-1 w.r.t. growth inhibitory  
 CC bioactivity. In contrast, however, ET-2 is apparently not capable of  
 CC eliciting the growth stimulatory activity characteristic of ET-1 and,  
 CC in fact, antagonises this ET-1 activity.  
 CC See also AAR14338-40, AAR14952-53, AAR14328-9 and AAR15315-20.

XX SQ Sequence 589 AA;

Query Match 68.9%; Score 51; DB 12; Length 589;

Best Local Similarity 66.7%; Pred. No. 1.3;

Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 3 RGTCLRRREAPR 14

:|||||::||

Db 564 RGTCLRRKTKPR 575

RESULT 13

AAU48409

ID AAR48409 standard; Protein; 101 AA.

XX AC AAR48409;

XX DT 27-FEB-2002 (first entry)

XX Propionibacterium acnes immunogenic protein #9305.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
 uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
 inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
 dermatological; osteopathic; neuroprotectant.

XX OS Propionibacterium acnes.

XX PN W0200181581-A2.

XX PD 01-NOV-2001.

XX PE 20-APR-2001; 2001WO-US12865.

XX PR 21-APR-2000; 2000US-199047P.

XX PR 02-JUN-2000; 2000US-208841P.

XX PR 07-JUL-2000; 2000US-216747P.

XX PA (CORI-) CORIXA CORP.

XX PI Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

XX PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

XX DR N-PSDB; AAS59542.

XX Propionibacterium acnes polypeptides and nucleic acids useful for  
 PT vaccinating against and diagnosing infections, especially useful for  
 PT treating acne vulgaris -

PS Example 1; SEQ ID NO 9604; 1069pp; English.

CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
CC polypeptides. The proteins and their associated DNA sequences are used in  
CC the treatment, prevention and diagnosis of medical conditions caused by  
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.  
CC P. acnes is also involved in infections of bone, joints and the central  
CC nervous system, however it is particularly involved in the inflammatory  
CC lesions associated with acne vulgaris. A method for detecting the  
CC presence or absence of P. acnes in a patient comprises contacting a  
CC sample with a binding agent that binds to the proteins of the invention  
CC and determining the amount of bound protein in the sample. The  
CC polypeptides may be used as antigens in the production of antibodies  
CC specific for P. acnes proteins. These antibodies can be used to  
CC downregulate expression and activity of P. acnes polypeptides and  
CC therefore treat P. acnes infections. The antibodies may also be used as  
CC diagnostic agents for determining P. acnes presence, for example, by  
CC enzyme linked immunosorbent assay (ELISA).  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 101 AA;

Query Match 62.2%; Score 46; DB 22; Length 101;  
Best Local Similarity 64.3%; Pred. No. 1.8;  
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 ARRGTKCLRREAPR 14  
DB 37 ARRGSCGRRERPR 50

RESULT 14

AAAR41885  
ID AAR41885 standard; peptide; 56 AA.

AC AAR41885;

DT 22-APR-1994 (first entry)

DE Granulin E.

KW Granulin; keratinocytes; wound healing; inhibition; peptide;  
granulocytes; leucocytes.

OS Homo sapiens.

PN WO9315195-A.

PD 05-AUG-1993.

PF 28-FEB-1992; 92WO-CA00089.

PR 03-FEB-1992; 92US-0829233.

PA (SOLO/) SOLOMON S.

PI Solomon S;

DR WPI; 1993-320328/40.

PT New cystine rich granulin peptide(s) from leucocyte(s) - are  
keratinocyte inhibitors useful topically for wound healing

PS Claim 8; Page 33; 53pp; English.

CC The granulin inhibits keratinocytes and is useful in formulations  
for promoting the healing of wounds. This peptide was designated  
granulin E.

XX Sequence 56 AA;

Query Match 58.1%; Score 43; DB 14; Length 56;  
Best Local Similarity 100.0%; Pred. No. 3.4;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ARRGTKCL 8  
DB 49 ARRGTKCL 56

RESULT 15

AAAY29199  
ID AAY29199 standard; Protein; 113 AA.

AC AAY29199;

DT 25-OCT-1999 (first entry)

DE Amino acid sequence of a virulence factor encoded by ORF26757c.

KW Human pathogen; virulence polypeptide; virulence factor;

pathogenic infection; Pseudomonas aeruginosa infection.

OS Pseudomonas aeruginosa.

PN WO9927129-A1.

PD 03-JUN-1999.

PF 25-NOV-1998; 98WO-US25247.

PR 25-NOV-1997; 97US-0066517.

PA (GEO) GEN HOSPITAL CORP.

PI Ausubel F, Cao H, Drenkard E, Goodman HM, Mahajan-Miklos S;

Rahme LG, Tan M, Tsongalis J;

WPI; 1999-357851/30.

PT Virulence factors useful in developing disease treatments

PS Disclosure; Fig 3; 228pp; English.

CC The present sequence represents a Pseudomonas aeruginosa polypeptide  
sequence. P. aeruginosa is an opportunistic human pathogen present in  
CC soil water and plants. The specification describes virulence polypeptides  
CC and nucleic acid sequence encoding such polypeptides. These sequences  
CC can be used to identify a compound which is capable of decreasing the  
CC expression of a pathogenic virulence factor. Compounds that inhibit  
CC the expression or activity of virulence factor polypeptides can be  
CC used to treat pathogenic infections, especially where the infection  
CC is a P. aeruginosa infection.  
CC note: the sequences given in the specification were poorly legible; and  
CC in some instances assumptions were made as to the identity of the  
CC residue; it is therefore possible that the sequence given below is  
CC not entirely correct.

SQ Sequence 113 AA;

Query Match 55.4%; Score 41; DB 20; Length 113;  
Best Local Similarity 61.5%; Pred. No. 15;  
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 RRGTKCLRREAPR 14

DB 94 RRGPRCRRRARR 106

Search completed: July 8, 2003, 16:24:17  
Job time : 36 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2003, 16:21:00 ; Search time 8.61539 Seconds  
(without alignments)  
47.812 Million cell updates/sec

Title: US-09-824-647-7  
Perfect score: 74  
Sequence: 1 ARGTGKCLRREAPR 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/PTCUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	74	100.0	14	4	US-08-991-862-7
2	74	100.0	593	4	US-08-991-862-17
3	68	91.9	593	1	US-07-668-648-4
4	68	91.9	593	2	US-08-429-998-4
5	68	91.9	593	2	US-08-431-333-4
6	68	91.9	593	5	PTC-US91-02321-4
7	53	71.6	14	4	US-08-991-862-5
8	53	71.6	589	1	US-07-668-648-6
9	53	71.6	589	2	US-08-429-998-6
10	53	71.6	589	2	US-08-431-333-6
11	53	71.6	589	4	US-08-991-862-2
12	53	71.6	589	5	PTC-US91-02321-6
13	51	68.9	589	1	US-07-668-648-2
14	51	68.9	589	2	US-08-429-998-2
15	51	68.9	589	2	US-08-431-333-2
16	51	68.9	589	5	PTC-US91-02321-2
17	41	55.4	113	4	US-09-199-637A-307
18	41	55.4	301	4	US-09-230-421-2
19	40	54.1	422	2	US-08-712-072C-2
20	40	54.1	1156	3	US-08-996-083-1
21	40	54.1	1156	4	US-09-429-516-1
22	39	52.7	278	4	US-09-429-516-3
23	39	52.7	278	4	US-09-724-864-52
24	39	52.7	1121	4	US-09-171-461-28
25	38	51.4	368	4	US-09-697-367-24
26	38	51.4	1068	2	US-08-429-054A-11
27	38	51.4	1068	2	US-08-718-777-7

Sequence 7, Appli  
Sequence 5, Appli  
Sequence 8, Appli  
Sequence 42, Appli  
Sequence 7, Appli  
Sequence 7, Appli  
Sequence 7, Appli  
Sequence 7, Appli  
Sequence 7, Appli  
Sequence 7, Appli  
Sequence 32, Appli  
Sequence 5, Appli  
Sequence 5, Appli  
Sequence 5, Appli  
Sequence 4, Appli  
Sequence 4, Appli  
Sequence 4, Appli

28 38 51.4 1068 3 US-09-051-341-7  
29 36 48.6 348 4 US-08-853-948B-5  
30 36 48.6 1257 3 US-08-611-729A-8  
31 36 48.6 2539 4 US-09-413-814-42  
32 41 32 47.3 41 1 US-08-112-208C-7  
33 35 47.3 41 1 US-08-248-819A-7  
34 35 47.3 41 2 US-08-337-646A-7  
35 35 47.3 41 2 US-08-856-531-7  
36 35 47.3 41 2 US-08-856-034-7  
37 35 47.3 41 4 US-08-927-326-7  
38 35 47.3 80 4 US-09-370-838-32  
39 35 47.3 263 3 US-09-035-706-5  
40 35 47.3 263 3 US-08-955-841-5  
41 35 47.3 263 4 US-09-390-425-5  
42 35 47.3 263 4 US-09-566-906-5  
43 35 47.3 441 2 US-08-491-835-4  
44 35 47.3 441 3 US-08-946-092A-4  
45 35 47.3 441 4 US-09-172-062-4

ALIGNMENTS

RESULT 1  
US-08-991-862-7  
; Sequence 7, Application US/08991862  
; Patent No. 6309826  
; GENERAL INFORMATION:  
; APPLICANT: Seriero, Ginette  
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
; FILE REFERENCE: 29996.488/P001-A  
; CURRENT FILING DATE: 1998-08-17  
; EARLIER APPLICATION NUMBER: US/08/991,862  
; EARLIER FILING DATE: 1997-05-23  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Human granulin  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(14)  
; OTHER INFORMATION: Internal peptide of human GP88 used to develop  
; neutralizing anti-human GP88 monoclonal antibody.  
US-08-991-862-7

Query Match 100.0%; Score 74; DB 4; Length 14;  
Best Local Similarity 100.0%; Pred. No. 1e-06; Mismatches 0; Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARGTGKCLRREAPR 14  
|||||  
DB 1 ARGTGKCLRREAPR 14

RESULT 2  
US-08-991-862-17  
; Sequence 17, Application US/08991862  
; Patent No. 6309826  
; GENERAL INFORMATION:  
; APPLICANT: Seriero, Ginette  
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
; FILE REFERENCE: 29996.488/P001-A  
; CURRENT FILING DATE: 1998-08-17  
; EARLIER APPLICATION NUMBER: US/08/991,862  
; EARLIER FILING DATE: 1997-05-23  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 17  
; LENGTH: 593

; TYPE: PRT  
; ORGANISM: Human GP88 CDNA  
US-08-991-862-17

Query Match 100.0%; Score 74; DB 4; Length 593;  
Best Local Similarity 100.0%; Pred. No. 4.3e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AARGTKCLRREAPR 14  
| | | | | | | | | | | | | | | | | |  
Db 566 AARGTKCLRREAPR 579

RESULT 3  
US-07-668-648-4  
; Sequence 4, Application US/07668648  
; Patent No. 5416192  
; GENERAL INFORMATION:  
; APPLICANT: Shoyab, Mohammed  
; APPLICANT: PLOWMAN, Gregory D.  
; TITLE OF INVENTION: EPITHELIALS: NOVEL CYSTEINE-RICH GROWTH  
; TITLE OF INVENTION: MODULATING PROTEINS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/668,648  
; FILING DATE: 19910819  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mistrock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 5624-161-999  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)790-9090  
; TELEFAX: (212) 869-9741  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 593 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-668-648-4

Query Match 91.9%; Score 68; DB 1; Length 593;  
Best Local Similarity 92.9%; Pred. No. 0.00046;  
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AARGTKCLRREAPR 14  
| | | | | | | | | | | | | | | | | |  
Db 566 AARGTKCLRREAPR 579

RESULT 4  
US-08-429-998-4  
; Sequence 4, Application US/08429998  
; Patent No. 5865961  
; GENERAL INFORMATION:  
; APPLICANT: Shoyab, Mohammed  
; APPLICANT: PLOWMAN, Gregory D.  
; TITLE OF INVENTION: EPITHELIALS: NOVEL CYSTEINE-RICH GROWTH  
; TITLE OF INVENTION: MODULATING PROTEINS  
; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/429,998  
; FILING DATE: 27-APR-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/668,648  
; FILING DATE: 13-MAR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mistrock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 5624-161-999  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)790-9090  
; TELEFAX: (212) 869-9741  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 593 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-429-998-4

Query Match 91.9%; Score 68; DB 2; Length 593;  
Best Local Similarity 92.9%; Pred. No. 0.00046;  
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AARGTKCLRREAPR 14  
| | | | | | | | | | | | | | | | | |  
Db 566 AARGTKCLRREAPR 579

RESULT 5  
US-08-431-333-4  
; Sequence 4, Application US/08431333  
; Patent No. 5965723  
; GENERAL INFORMATION:  
; APPLICANT: Shoyab, Mohammed  
; APPLICANT: PLOWMAN, Gregory D.  
; TITLE OF INVENTION: EPITHELIALS: NOVEL CYSTEINE-RICH GROWTH  
; TITLE OF INVENTION: MODULATING PROTEINS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/431,333  
; FILING DATE: 27-APR-1995  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/668,648  
; FILING DATE: 13-MAR-1991  
; ATTORNEY/AGENT INFORMATION:

Query Match 91.9%; Score 68; DB 1; Length 593;  
Best Local Similarity 92.9%; Pred. No. 0.00046;  
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AARGTKCLRREAPR 14  
| | | | | | | | | | | | | | | | | |  
Db 566 AARGTKCLRREAPR 579

RESULT 4  
US-08-429-998-4  
; Sequence 4, Application US/08429998  
; Patent No. 5865961  
; GENERAL INFORMATION:  
; APPLICANT: Shoyab, Mohammed  
; APPLICANT: PLOWMAN, Gregory D.  
; TITLE OF INVENTION: EPITHELIALS: NOVEL CYSTEINE-RICH GROWTH  
; TITLE OF INVENTION: MODULATING PROTEINS  
; NUMBER OF SEQUENCES: 12

NAME: Mistrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)790-9090  
TELEFAX: (212) 869-9741  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 593 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-431-333-4

Query Match 91.9%; Score 68; DB 2; Length 593;  
Best Local Similarity 92.9%; Pred. No. 0.00046;  
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AARGTKCLRREAPR 14  
| | | | | | | | | | | | | | | |  
Db 566 AARGTKCLRREAPR 579

## RESULT 6

PCT-US91-02321-4  
Sequence 4, Application PC/TUS9102321  
GENERAL INFORMATION:  
APPLICANT: Shoyab, Mohammed  
TITLE OF INVENTION: EPIITHELINS: NOVEL CYSTEINE-RICH GROWTH  
TITLE OF INVENTION: MODULATING PROTEINS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Bristol-Myers Squibb Company  
STREET: 3005 First Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98121  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/02321  
FILING DATE: 19910403  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Poor, Brian W.  
REGISTRATION NUMBER: 32,928  
REFERENCE/DOCKET NUMBER: ON0071A-PC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)728-4800  
TELEFAX: (206)448-4775  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 593 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US91-02321-4

Query Match 91.9%; Score 68; DB 5; Length 593;  
Best Local Similarity 92.9%; Pred. No. 0.00046;  
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AARGTKCLRREAPR 14  
| | | | | | | | | | | | | | | |  
Db 566 AARGTKCLRREAPR 579

## RESULT 7

US-08-991-862-5  
Sequence 5, Application US/08991862  
Patent No. 6309826  
GENERAL INFORMATION:  
APPLICANT: Serrero, Ginette  
TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
FILE REFERENCE: 29996.488/P001-A  
CURRENT APPLICATION NUMBER: US/08/991.862  
CURRENT FILING DATE: 1998-08-17  
EARLIER APPLICATION NUMBER: 08/863,862  
EARLIER FILING DATE: 1997-03-23  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 5  
LENGTH: 14  
TYPE: PRT  
ORGANISM: mouse granulin  
FEATURE:  
NAME/KEY: PEPTIDE  
LOCATION: (1)..(14)  
OTHER INFORMATION: Internal peptide of mouse GP88 used to raise the  
OTHER INFORMATION: antisera against the GP88 used in the  
OTHER INFORMATION: immunoaffinity step.  
US-08-991-862-5

Query Match 71.6%; Score 53; DB 4; Length 14;  
Best Local Similarity 75.0%; Pred. No. 0.0043;  
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RGTGKCLRREAPR 14  
| | | | | | | | | | | | | | | |  
Db 3 RGTGKCLRREAPR 14

## RESULT 8

US-07-668-648-6  
Sequence 6, Application US/07668648  
Patent No. 5416192  
GENERAL INFORMATION:  
APPLICANT: Shoyab, Mohammed  
APPLICANT: Plowman, Gregory D.  
TITLE OF INVENTION: EPIITHELINS: NOVEL CYSTEINE-RICH GROWTH  
TITLE OF INVENTION: MODULATING PROTEINS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/668,648  
FILING DATE: 19910819  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)790-9090  
TELEFAX: (212) 869-9741  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 589 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein



US-07-668-648-6

Query Match 71.6%; Score 53; DB 1; Length 589;  
Best Local Similarity 75.0%; Pred. No. 0.18;  
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RGTKCLRREAPR 14  
|||||: ||  
Db 564 RGTKCLRKKIPR 575

RESULT 9

US-08-429-998-6  
Sequence 6, Application US/08429998  
Patent No. 5885961  
GENERAL INFORMATION:  
APPLICANT: Shoyab, Mohammed  
APPLICANT: Plowman, Gregory D.  
TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH  
MODULATING PROTEINS  
TITLE OF INVENTION: MODULATING PROTEINS  
NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/429,998  
FILING DATE: 27-APR-1995

CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/668,648  
FILING DATE: 13-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELEPHONE: (212)790-9090  
TELEFAX: (212) 869-9741  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 589 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-429-998-6

Query Match 71.6%; Score 53; DB 2; Length 589;  
Best Local Similarity 75.0%; Pred. No. 0.18;  
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RGTKCLRREAPR 14  
|||||: ||  
Db 564 RGTKCLRKKIPR 575

RESULT 10

US-08-431-333-6  
Sequence 6, Application US/08431333  
Patent No. 5965723  
GENERAL INFORMATION:  
APPLICANT: Shoyab, Mohammed  
APPLICANT: Plowman, Gregory D.  
TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH  
MODULATING PROTEINS  
TITLE OF INVENTION: MODULATING PROTEINS

NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/431,333  
FILING DATE: 27-APR-1995  
CLASSIFICATION: 536

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/668,648  
FILING DATE: 13-MAR-1991

ATTORNEY/AGENT INFORMATION:  
NAME: Mistrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELEPHONE: (212)790-9090  
TELEFAX: (212) 869-9741  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 589 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-431-333-6

Query Match 71.6%; Score 53; DB 2; Length 589;  
Best Local Similarity 75.0%; Pred. No. 0.18;  
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RGTKCLRREAPR 14  
|||||: ||  
Db 564 RGTKCLRKKIPR 575

RESULT 11

US-08-991-862-2  
Sequence 2, Application US/08991862  
Patent No. 6309826  
GENERAL INFORMATION:  
APPLICANT: Serrero, Ginette  
TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
FILE REFERENCE: 29996.488/P001-A  
CURRENT APPLICATION NUMBER: US/08/991,862  
CURRENT FILING DATE: 1998-08-17  
EARLIER APPLICATION NUMBER: 08/863,862  
EARLIER FILING DATE: 1997-05-23  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 2  
LENGTH: 589  
TYPE: PRT  
ORGANISM: Mouse epithelin/granulin

US-08-991-862-2

Query Match 71.6%; Score 53; DB 4; Length 589;  
Best Local Similarity 75.0%; Pred. No. 0.18;  
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RGTKCLRREAPR 14  
|||||: ||  
Db 564 RGTKCLRKKIPR 575

RESULT 12  
PCT-US91-02321-6  
; Sequence 6, Application PC/TUS9102321  
; GENERAL INFORMATION:  
; APPLICANT: Shoyab, Mohammed  
; APPLICANT: Plozman, Gregory D.  
; TITLE OF INVENTION: NOVEL CYSTEINE-RICH GROWTH  
; TITLE OF INVENTION: MODULATING PROTEINS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Bristol-Myers Squibb Company  
; STREET: 3005 First Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98121  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US91/02321  
; FILING DATE: 19910403  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Poor, Brian W.  
; REGISTRATION NUMBER: 32,928  
; REFERENCE/DOCKET NUMBER: OM0071A-PC  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206)728-4800  
; TELEFAX: (206)448-4775  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 589 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US91-02321-6  
Query Match 71.6%; Score 53; DB 5; Length 589;  
Best Local Similarity 75.0%; Pred. No. 0.18;  
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 3 RGTKCLRRAPR 14  
DB 564 RGTKCLRRKIPR 575  
RESULT 13  
US-07-668-648-2  
; Sequence 2, Application US/07668648  
; Patent No. 5416192  
; GENERAL INFORMATION:  
; APPLICANT: Shoyab, Mohammed  
; APPLICANT: Plozman, Gregory D.  
; TITLE OF INVENTION: NOVEL CYSTEINE-RICH GROWTH  
; TITLE OF INVENTION: MODULATING PROTEINS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/668,648

; FILING DATE: 19910819  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Misrock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 5624-161-999  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)790-9090  
; TELEFAX: (212) 869-9741  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 589 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-07-668-648-2  
Query Match 68.9%; Score 51; DB 1; Length 589;  
Best Local Similarity 66.7%; Pred. No. 0.39;  
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 3 RGTKCLRRAPR 14  
DB 564 RGTKCLRRKIPR 575  
RESULT 14  
US-08-429-998-2  
; Sequence 2, Application US/08429998  
; Patent No. 5885961  
; GENERAL INFORMATION:  
; APPLICANT: Shoyab, Mohammed  
; APPLICANT: Plozman, Gregory D.  
; TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH  
; TITLE OF INVENTION: MODULATING PROTEINS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/429,998  
; FILING DATE: 27-APR-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/668,648  
; FILING DATE: 13-MAR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Misrock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 5624-161-999  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)790-9090  
; TELEFAX: (212) 869-9741  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 589 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-429-998-2  
Query Match 68.9%; Score 51; DB 2; Length 589;  
Best Local Similarity 66.7%; Pred. No. 0.39;  
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 RGTKCLRRREAPR 14  
:|||||: ||  
Db 564 RGTKCLRRKTPR 575

RESULT 15  
US-08-431-333-2  
: Sequence 2, Application US/08431333  
: Patent No. 5965723  
: GENERAL INFORMATION:  
: APPLICANT: Shoyab, Mohammed  
: APPLICANT: Plozman, Gregory D.  
: TITLE OF INVENTION: EPIHELINS: NOVEL CYSTEINE-RICH GROWTH  
: TITLE OF INVENTION: MODULATING PROTEINS  
: NUMBER OF SEQUENCES: 12  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Pennie & Edmonds  
: STREET: 1155 Avenue of the Americas  
: CITY: New York  
: STATE: New York  
: COUNTRY: USA  
: ZIP: 10036  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patent In Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/431,333  
: FILING DATE: 27-APR-1995  
: CLASSIFICATION: 536  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 07/668,648  
: FILING DATE: 13-MAR-1991  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Misrock, S. Leslie  
: REGISTRATION NUMBER: 18,872  
: REFERENCE/DOCKET NUMBER: 5624-161-999  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (212)790-9090  
: TELEFAX: (212) 869-9741  
: INFORMATION FOR SEQ ID NO: 2:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 589 amino acids  
: TYPE: amino acid  
: TOPOLOGY: linear  
: MOLECULE TYPE: protein  
US-08-431-333-2

Query Match 68.9%; Score 51; DB 2; Length 589;  
Best Local Similarity 66.7%; Pred. No. 0.39;  
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 RGTKCLRRREAPR 14  
:|||||: ||  
Db 564 RGTKCLRRKTPR 575

Search completed: July 8, 2003, 16:30:21  
Job time : 9.61539 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2003, 16:24:26 ; Search time 13.641 Seconds  
(without alignments)  
119.483 Million cell updates/sec

Title: US-09-824-647-7

Perfect score: 74

Sequence: 1 ARRGTKLRREAPR 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 116419773 residues

Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	74	100.0	14	9 US-09-824-647-7	Sequence 7, Appli
2	74	100.0	14	9 US-10-218-509-7	Sequence 7, Appli
3	74	100.0	14	9 US-10-281-160-7	Sequence 7, Appli
4	74	100.0	14	10 US-09-813-156-7	Sequence 7, Appli
5	74	100.0	14	10 US-09-824-807-7	Sequence 7, Appli
6	74	100.0	593	9 US-09-824-647-17	Sequence 17, Appl
7	74	100.0	593	9 US-10-218-509-17	Sequence 17, Appl
8	74	100.0	593	9 US-10-281-160-17	Sequence 17, Appl
9	74	100.0	593	10 US-09-813-156-17	Sequence 17, Appl
10	74	100.0	593	10 US-09-824-807-17	Sequence 17, Appl
11	68	91.9	621	10 US-09-925-301-1416	Sequence 1416, Ap
12	53	71.6	14	9 US-09-824-647-5	Sequence 5, Appli
13	53	71.6	14	9 US-10-218-509-5	Sequence 5, Appli
14	53	71.6	14	9 US-10-281-160-5	Sequence 5, Appli
15	53	71.6	14	10 US-09-813-156-5	Sequence 5, Appli
16	53	71.6	14	10 US-09-824-807-5	Sequence 5, Appli
17	53	71.6	589	9 US-09-824-647-2	Sequence 2, Appli
18	53	71.6	589	9 US-10-218-509-2	Sequence 2, Appli
19	53	71.6	589	9 US-10-281-160-2	Sequence 2, Appli

20	53	71.6	589	10 US-09-813-156-2	Sequence 2, Appli
21	53	71.6	589	10 US-09-824-807-2	Sequence 2, Appli
22	41	55.4	113	9 US-09-975-719-307	Sequence 307, App
23	41	55.4	1083	9 US-10-217-700-11	Sequence 11, Appl
24	40	54.1	250	9 US-10-156-761-14524	Sequence 14524, A
25	40	54.1	1156	10 US-09-757-716-1	Sequence 1, Appli
26	39	52.7	255	9 US-09-866-050A-692	Sequence 692, App
27	39	52.7	275	9 US-10-269-353-15	Sequence 15, Appl
28	39	52.7	275	10 US-09-790-264-15	Sequence 15, Appl
29	39	52.7	1121	10 US-09-970-711-28	Sequence 28, Appl
30	38	51.4	119	9 US-10-103-313-525	Sequence 525, App
31	38	51.4	137	9 US-10-103-313-397	Sequence 397, App
32	38	51.4	308	9 US-10-121-988-97	Sequence 97, Appl
33	38	51.4	846	9 US-10-156-761-13739	Sequence 13739, A
34	38	51.4	1068	9 US-10-217-700-8	Sequence 8, Appli
35	37	50.0	717	10 US-09-925-300-1299	Sequence 1299, Ap
36	37	50.0	903	9 US-09-746-783-142	Sequence 142, App
37	37	50.0	1049	9 US-10-217-700-10	Sequence 10, Appl
38	37	50.0	1081	9 US-10-217-700-4	Sequence 4, Appli
39	37	50.0	1084	9 US-10-217-700-9	Sequence 9, Appli
40	37	50.0	1087	10 US-09-918-909-24	Sequence 24, Appl
41	36.5	49.3	186	10 US-09-789-561-160	Sequence 160, App
42	36	48.6	51	9 US-09-764-891-4995	Sequence 4995, Ap
43	36	48.6	143	9 US-09-764-891-4973	Sequence 4973, Ap
44	36	48.6	145	9 US-09-764-891-5104	Sequence 5104, Ap
45	36	48.6	148	9 US-10-226-489-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1  
US-09-824-647-7  
; Sequence 7, Application US/09824647  
; Publication No. US20020183270A1  
; GENERAL INFORMATION:  
; APPLICANT: Serrero, Ginette  
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
; FILE REFERENCE: 29996.488/P001-A  
; CURRENT APPLICATION NUMBER: US/09/824,647  
; CURRENT FILING DATE: 2001-04-04  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/991,862  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/863,862  
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-05-23  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Human granulin  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(14)  
; OTHER INFORMATION: Internal peptide of human GP88 used to develop  
; OTHER INFORMATION: neutralizing anti-human GP88 monoclonal antibody.  
US-09-824-647-7

Query Match 100.0%; Score 74; DB 9; Length 14;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 ARRGTKLRREAPR 14  
Db 1 ARRGTKLRREAPR 14

RESULT 2  
US-10-218-509-7  
; Sequence 7, Application US/10218509  
; Publication No. US20030092661A1  
; GENERAL INFORMATION:  
; APPLICANT: Serrero, Ginette

;; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
;; FILE REFERENCE: Z9996.488/P001-A  
;; CURRENT APPLICATION NUMBER: US/10/218,509  
;; CURRENT FILING DATE: 2002-08-15  
;; PRIOR APPLICATION NUMBER: 08/991,862  
;; PRIOR FILING DATE: 1998-08-17  
;; PRIOR APPLICATION NUMBER: 08/863,862  
;; PRIOR FILING DATE: 1997-05-23  
;; NUMBER OF SEQ ID NOS: 17  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 7  
;; LENGTH: 14  
;; TYPE: PRT  
;; ORGANISM: Human granulin  
;; FEATURE:  
;; NAME/KEY: PEPTIDE  
;; LOCATION: (1)..(14)  
;; OTHER INFORMATION: Internal peptide of human GP88 used to develop  
;; OTHER INFORMATION: neutralizing anti-human GP88 monoclonal antibody.  
US-10-218-509-7

Query Match 100.0%; Score 74; DB 9; Length 14;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARRGTKCLRREAPR 14  
|||||  
DB 1 ARRGTKCLRREAPR 14

RESULT 3  
US-10-281-160-7  
;; Sequence 7, Application US/10281160  
;; Publication No. US20030108950A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Seriero, Ginette  
;; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
;; FILE REFERENCE: Z9996.488/P001-A  
;; CURRENT APPLICATION NUMBER: US/10/281,160  
;; CURRENT FILING DATE: 2002-10-28  
;; PRIOR APPLICATION NUMBER: US/08/991,862  
;; PRIOR FILING DATE: 1998-08-17  
;; PRIOR APPLICATION NUMBER: 08/863,862  
;; PRIOR FILING DATE: 1997-05-23  
;; NUMBER OF SEQ ID NOS: 17  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 7  
;; LENGTH: 14  
;; TYPE: PRT  
;; ORGANISM: Human granulin  
;; FEATURE:  
;; NAME/KEY: PEPTIDE  
;; LOCATION: (1)..(14)  
;; OTHER INFORMATION: Internal peptide of human GP88 used to develop  
;; OTHER INFORMATION: neutralizing anti-human GP88 monoclonal antibody.  
US-10-281-160-7

Query Match 100.0%; Score 74; DB 9; Length 14;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARRGTKCLRREAPR 14  
|||||  
DB 1 ARRGTKCLRREAPR 14

RESULT 4  
US-09-813-156-7  
;; Sequence 7, Application US/09813156  
;; Patent No. US20020061859A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Seriero, Ginette  
;; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS

;; FILE REFERENCE: Z9996.488/P001-A  
;; CURRENT APPLICATION NUMBER: US/09/813,156  
;; CURRENT FILING DATE: 2001-03-21  
;; PRIOR APPLICATION NUMBER: 08/991,862  
;; PRIOR FILING DATE: 1997-12-16  
;; PRIOR APPLICATION NUMBER: 08/863,862  
;; PRIOR FILING DATE: 1997-05-23  
;; NUMBER OF SEQ ID NOS: 17  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 7  
;; LENGTH: 14  
;; TYPE: PRT  
;; ORGANISM: Human granulin  
;; FEATURE:  
;; NAME/KEY: PEPTIDE  
;; LOCATION: (1)..(14)  
;; OTHER INFORMATION: Internal peptide of human GP88 used to develop  
;; OTHER INFORMATION: neutralizing anti-human GP88 monoclonal antibody.  
US-09-813-156-7

Query Match 100.0%; Score 74; DB 10; Length 14;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARRGTKCLRREAPR 14  
|||||  
DB 1 ARRGTKCLRREAPR 14

RESULT 5  
US-09-824-807-7  
;; Sequence 7, Application US/09824807  
;; Patent No. US20020094966A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Seriero, Ginette  
;; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
;; FILE REFERENCE: Z9996.488/P001-A  
;; CURRENT APPLICATION NUMBER: US/09/824,807  
;; CURRENT FILING DATE: 2001-04-04  
;; PRIOR APPLICATION NUMBER: 08/991,862  
;; PRIOR FILING DATE: 1997-12-16  
;; PRIOR APPLICATION NUMBER: 08/863,862  
;; PRIOR FILING DATE: 1997-05-23  
;; NUMBER OF SEQ ID NOS: 17  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 7  
;; LENGTH: 14  
;; TYPE: PRT  
;; ORGANISM: Human granulin  
;; FEATURE:  
;; NAME/KEY: PEPTIDE  
;; LOCATION: (1)..(14)  
;; OTHER INFORMATION: Internal peptide of human GP88 used to develop  
;; OTHER INFORMATION: neutralizing anti-human GP88 monoclonal antibody.  
US-09-824-807-7

Query Match 100.0%; Score 74; DB 10; Length 14;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARRGTKCLRREAPR 14  
|||||  
DB 1 ARRGTKCLRREAPR 14

RESULT 6  
US-09-824-647-17  
;; Sequence 17, Application US/09824647  
;; Publication No. US20020183270A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Seriero, Ginette  
;; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
;; FILE REFERENCE: Z9996.488/P001-A

```
; CURRENT APPLICATION NUMBER: US/09/824,647
; CURRENT FILING DATE: 2001-04-04
; PRIOR FILING DATE: EARLIER APPLICATION NUMBER: 08/991,862
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/863,862
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-05-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Human GP88 cDNA
US-09-824-647-17

Query Match      100.0%; Score 74; DB 9; Length 593;
Best Local Similarity 100.0%; Pred. No. 7.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ARRGTKCLRREAPR 14
        |||
Db      566 ARRGTKCLRREAPR 579

RESULT 7
US-10-218-509-17
; Sequence 17, Application US/10218509
; Publication No. US20030092661A1
; GENERAL INFORMATION:
; APPLICANT: Serrero, Ginette
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
; FILE REFERENCE: Z9996.488/P001-A
; CURRENT APPLICATION NUMBER: US/10/218,509
; CURRENT FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: 08/991,862
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 08/863,862
; PRIOR FILING DATE: 1997-05-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Human GP88 cDNA
US-10-218-509-17

Query Match      100.0%; Score 74; DB 9; Length 593;
Best Local Similarity 100.0%; Pred. No. 7.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ARRGTKCLRREAPR 14
        |||
Db      566 ARRGTKCLRREAPR 579

RESULT 8
US-10-281-160-17
; Sequence 17, Application US/10281160
; Publication No. US20030108950A1
; GENERAL INFORMATION:
; APPLICANT: Serrero, Ginette
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
; FILE REFERENCE: Z9996.488/P001-A
; CURRENT APPLICATION NUMBER: US/10/281,160
; CURRENT FILING DATE: 2002-10-28
; PRIOR APPLICATION NUMBER: US/08/991,862
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 08/863,862
; PRIOR FILING DATE: 1997-05-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Human GP88 cDNA
US-10-281-160-17

Query Match      100.0%; Score 74; DB 9; Length 593;
Best Local Similarity 100.0%; Pred. No. 7.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ARRGTKCLRREAPR 14
        |||
Db      566 ARRGTKCLRREAPR 579

RESULT 9
US-09-813-156-17
; Sequence 17, Application US/09813156
; Patent No. US20020061859A1
; GENERAL INFORMATION:
; APPLICANT: Serrero, Ginette
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
; FILE REFERENCE: Z9996.488/P001-A
; CURRENT APPLICATION NUMBER: US/09/813,156
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 08/991,862
; PRIOR FILING DATE: 1997-12-16
; PRIOR APPLICATION NUMBER: 08/863,862
; PRIOR FILING DATE: 1997-05-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Human GP88 cDNA
US-09-813-156-17

Query Match      100.0%; Score 74; DB 10; Length 593;
Best Local Similarity 100.0%; Pred. No. 7.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ARRGTKCLRREAPR 14
        |||
Db      566 ARRGTKCLRREAPR 579

RESULT 10
US-09-824-807-17
; Sequence 17, Application US/09824807
; Patent No. US20020094966A1
; GENERAL INFORMATION:
; APPLICANT: Serrero, Ginette
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
; FILE REFERENCE: Z9996.488/P001-A
; CURRENT APPLICATION NUMBER: US/09/824,807
; CURRENT FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 08/991,862
; PRIOR FILING DATE: 1997-12-16
; PRIOR APPLICATION NUMBER: 08/863,862
; PRIOR FILING DATE: 1997-05-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Human GP88 cDNA
US-09-824-807-17

Query Match      100.0%; Score 74; DB 10; Length 593;
Best Local Similarity 100.0%; Pred. No. 7.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ARRGTKCLRREAPR 14
        |||
Db      566 ARRGTKCLRREAPR 579
```

```
; ORGANISM: Human GP88 cDNA
US-10-281-160-17

Query Match      100.0%; Score 74; DB 9; Length 593;
Best Local Similarity 100.0%; Pred. No. 7.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ARRGTKCLRREAPR 14
        |||
Db      566 ARRGTKCLRREAPR 579

RESULT 9
US-09-813-156-17
; Sequence 17, Application US/09813156
; Patent No. US20020061859A1
; GENERAL INFORMATION:
; APPLICANT: Serrero, Ginette
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
; FILE REFERENCE: Z9996.488/P001-A
; CURRENT APPLICATION NUMBER: US/09/813,156
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 08/991,862
; PRIOR FILING DATE: 1997-12-16
; PRIOR APPLICATION NUMBER: 08/863,862
; PRIOR FILING DATE: 1997-05-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Human GP88 cDNA
US-09-813-156-17

Query Match      100.0%; Score 74; DB 10; Length 593;
Best Local Similarity 100.0%; Pred. No. 7.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ARRGTKCLRREAPR 14
        |||
Db      566 ARRGTKCLRREAPR 579

RESULT 10
US-09-824-807-17
; Sequence 17, Application US/09824807
; Patent No. US20020094966A1
; GENERAL INFORMATION:
; APPLICANT: Serrero, Ginette
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
; FILE REFERENCE: Z9996.488/P001-A
; CURRENT APPLICATION NUMBER: US/09/824,807
; CURRENT FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 08/991,862
; PRIOR FILING DATE: 1997-12-16
; PRIOR APPLICATION NUMBER: 08/863,862
; PRIOR FILING DATE: 1997-05-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Human GP88 cDNA
US-09-824-807-17

Query Match      100.0%; Score 74; DB 10; Length 593;
Best Local Similarity 100.0%; Pred. No. 7.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ARRGTKCLRREAPR 14
        |||
Db      566 ARRGTKCLRREAPR 579
```

## RESULT 11

US-09-925-301-1416  
 ; Sequence 1416, Application US/09925301  
 ; Patent No. US20020052308A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
 ; FILE REFERENCE: PA106  
 ; CURRENT APPLICATION NUMBER: US/09/925,301  
 ; CURRENT FILING DATE: 2001-08-10  
 ; PRIOR APPLICATION NUMBER: PCT/US00/05982  
 ; PRIOR FILING DATE: 2000-03-08  
 ; PRIOR APPLICATION NUMBER: 60/134,270  
 ; PRIOR FILING DATE: 1999-03-12  
 ; NUMBER OF SEQ ID NOS: 1694  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 1416  
 ; LENGTH: 621  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-925-301-1416

Query Match 91.9%; Score 68; DB 10; Length 621;  
 Best Local Similarity 92.9%; Pred. No. 0.00092;  
 Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ARGTGKCLRREAPR 14  
 | | | | | | | | | | | | | | | |  
 Db 594 AARGTKCLRREAPR 607

## RESULT 12

US-09-824-647-5  
 ; Sequence 5, Application US/09824647  
 ; Publication No. US20020183270A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Serrero, Ginette  
 ; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
 ; FILE REFERENCE: 29996.488/P001-A  
 ; CURRENT APPLICATION NUMBER: US/09/824,647  
 ; CURRENT FILING DATE: 2001-04-04  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/991,862  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-17  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/863,862  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1997-05-23  
 ; NUMBER OF SEQ ID NOS: 17  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 5  
 ; LENGTH: 14  
 ; TYPE: PRT  
 ; ORGANISM: mouse granuln  
 ; FEATURE:  
 ; NAME/KEY: PEPTIDE  
 ; LOCATION: (1)..(14)  
 ; OTHER INFORMATION: Internal peptide of mouse GP88 used to raise the  
 ; OTHER INFORMATION: antiserum against the GP88 used in the  
 ; OTHER INFORMATION: immunoaffinity step.  
 US-09-824-647-5

Query Match 71.6%; Score 53; DB 9; Length 14;  
 Best Local Similarity 75.0%; Pred. No. 0.0072;  
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RGTGKCLRREAPR 14  
 | | | | | | | | | | | | | | | |  
 Db 3 RGTGKCLRREAPR 14

## RESULT 13

US-10-218-509-5  
 ; Sequence 5, Application US/10218509  
 ; Publication No. US20030092661A1

; GENERAL INFORMATION:  
 ; APPLICANT: Serrero, Ginette  
 ; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
 ; FILE REFERENCE: 29996.488/P001-A  
 ; CURRENT APPLICATION NUMBER: US/10/218,509  
 ; CURRENT FILING DATE: 2002-08-15  
 ; PRIOR APPLICATION NUMBER: 08/991,862  
 ; PRIOR FILING DATE: 1998-08-17  
 ; PRIOR APPLICATION NUMBER: 08/863,862  
 ; PRIOR FILING DATE: 1997-05-23  
 ; NUMBER OF SEQ ID NOS: 17  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 5  
 ; LENGTH: 14  
 ; TYPE: PRT  
 ; ORGANISM: mouse granuln  
 ; FEATURE:  
 ; NAME/KEY: PEPTIDE  
 ; LOCATION: (1)..(14)  
 ; OTHER INFORMATION: Internal peptide of mouse GP88 used to raise the  
 ; OTHER INFORMATION: antiserum against the GP88 used in the  
 ; OTHER INFORMATION: immunoaffinity step.  
 US-10-218-509-5

Query Match 71.6%; Score 53; DB 9; Length 14;  
 Best Local Similarity 75.0%; Pred. No. 0.0072;  
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RGTGKCLRREAPR 14  
 | | | | | | | | | | | | | | | |  
 Db 3 RGTGKCLRREAPR 14

## RESULT 14

US-10-281-160-5  
 ; Sequence 5, Application US/10281160  
 ; Publication No. US20030108950A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Serrero, Ginette  
 ; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
 ; FILE REFERENCE: 29996.488/P001-A  
 ; CURRENT APPLICATION NUMBER: US/10/281,160  
 ; CURRENT FILING DATE: 2002-10-28  
 ; PRIOR APPLICATION NUMBER: US/08/991,862  
 ; PRIOR FILING DATE: 1998-08-17  
 ; PRIOR APPLICATION NUMBER: 08/863,862  
 ; PRIOR FILING DATE: 1997-05-23  
 ; NUMBER OF SEQ ID NOS: 17  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 5  
 ; LENGTH: 14  
 ; TYPE: PRT  
 ; ORGANISM: mouse granuln  
 ; FEATURE:  
 ; NAME/KEY: PEPTIDE  
 ; LOCATION: (1)..(14)  
 ; OTHER INFORMATION: Internal peptide of mouse GP88 used to raise the  
 ; OTHER INFORMATION: antiserum against the GP88 used in the  
 ; OTHER INFORMATION: immunoaffinity step.  
 US-10-281-160-5

Query Match 71.6%; Score 53; DB 9; Length 14;  
 Best Local Similarity 75.0%; Pred. No. 0.0072;  
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RGTGKCLRREAPR 14  
 | | | | | | | | | | | | | | | |  
 Db 3 RGTGKCLRREAPR 14

## RESULT 15

US-09-813-156-5  
 ; Sequence 5, Application US/09813156

; Patent No. US20020061859A1  
; GENERAL INFORMATION:  
; APPLICANT: Seriero, Ginette  
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS  
; FILE REFERENCE: Z9996.488/P001-A  
; CURRENT APPLICATION NUMBER: US/09/813,156  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 08/991,862  
; PRIOR FILING DATE: 1997-12-16  
; PRIOR APPLICATION NUMBER: 08/863,862  
; PRIOR FILING DATE: 1997-05-23  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: mouse granulin  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(14)  
; OTHER INFORMATION: Internal peptide of mouse GP88 used to raise the  
; OTHER INFORMATION: antisera against the GP88 used in the  
; OTHER INFORMATION: immunoaffinity step.  
US-09-813-156-5

Query Match 71.6%; Score 53; DB 10; Length 14;  
Best Local Similarity 75.0%; Pred. No. 0.0072;  
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RGTCLRRREAPR 14  
| | | | | | | | | | | | | | | |  
Db 3 RGTCLRRKKIPR 14

Search completed: July 8, 2003, 16:31:42  
Job time : 13.641 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein : protein search, using sw model

Run on: July 8, 2003, 16:19:45 ; Search time 14.1795 Seconds  
(without alignments)  
94.918 Million cell updates/sec

Title: US-09-824-647-7

Perfect score: 74

Sequence: 1 ARRTKCLREAPR 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_73:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	68	91.9	593	1 GYHU	granulin precursor
2	53	71.6	589	2 C38128	epithelin/granulin
3	51	68.9	589	2 B38128	epithelin/granulin
4	43	58.1	365	2 A75577	nicotinate-nucleot
5	41	55.4	327	2 T49514	hypothetical prote
6	41	55.4	422	2 T49513	gastric mucin rela
7	41	55.4	1083	2 T04062	sucrose-phosphate
8	41	55.4	1324	2 S06187	RNA2 polyprotein -
9	40	54.1	327	2 T19782	hypothetical prote
10	40	54.1	378	2 G70766	hypothetical prote
11	40	54.1	500	2 T50868	exopolysphatase
12	40	54.1	831	2 H84368	MCM / cell divisio
13	39	52.7	118	2 T08723	hypothetical prote
14	39	52.7	147	2 D70775	hypothetical prote
15	38	51.4	187	2 C75558	acetyl-CoA carboxy
16	38	51.4	1068	1 JQ1329	sucrose-phosphate
17	37	50.0	157	2 T02664	allergen - rice
18	37	50.0	246	2 AG2138	hypothetical prote
19	37	50.0	288	2 T24824	hypothetical prote
20	37	50.0	376	2 C75580	adenine deaminase-
21	37	50.0	379	2 AF2409	mannosyl transfera
22	37	50.0	598	2 A87257	methyl-accepting c
23	37	50.0	869	2 T22422	hypothetical prote
24	37	50.0	1047	2 T51800	sucrose-phosphate
25	37	50.0	1049	2 JCA783	sucrose-phosphate
26	37	50.0	1064	2 F86182	hypothetical prote
27	37	50.0	1081	2 T09837	sucrose-phosphate
28	37	50.0	1084	2 T04103	sucrose-phosphate
29	36.5	49.3	123	2 E72680	hypothetical prote

#### ALIGNMENTS

##### RESULT 1

GYHU

granulin precursor [validated] - human

N:Alternate names: epithelin

N:Contains: granulin A; granulin B; granulin C; granulin D; granulin E; granulin

C:Species: Homo sapiens (man)

C>Date: 30-Sep-1992 #sequence\_revision 03-May-1996 #text\_change 08-Dec-2000

C:Accession: JCI1284; A38128; A38118; A36698; B36698; C36698; D36698; A56873

R:Bhandari, V.; Bateman, A.

Biochem. Biophys. Res. Commun. 188, 57-63, 1992

A:Title: Structure and chromosomal location of the human granulin gene.

A:Reference number: JCI1284; MUID:93038704; PMID:1417868

A:Accession: JCI1284

A:Molecule type: DNA

A:Residues: 1-593 <BHA>

R:Plowman, G.D.; Green, J.M.; Neubauer, M.G.; Buckley, S.D.; McDonald, V.L.; Tod:

J. Biol. Chem. 267, 13073-13078, 1992

A:Title: The epithelin precursor encodes two proteins with opposing activities o

A:Reference number: A38128; MUID:92317004; PMID:1618805

A:Accession: A38128

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-593 <PLO>

A:Cross-references: GB:X62320; NID:931192; PIDN:CAA44196.1; PID:931193

R:Bhandari, V.; Palfree, R.G.E.; Bateman, A.

Proc. Natl. Acad. Sci. U.S.A. 89, 1715-1719, 1992

A:Title: Isolation and sequence of the granulin precursor cDNA from human bone m

A:Reference number: A38118; MUID:92179253; PMID:1542665

A:Accession: A38118

A:Molecule type: mRNA

A:Residues: 1-406, 'R', 408-433, 'G', 435-453, 'G', 455-459, 'Q', 461-546, 'A', 548-566, 'R

A:Cross-references: GB:M75161; NID:g183612; PIDN:AAA58617.1; PID:g183613

A:Note: this sequence has been revised in reference JCI1284

R:Bateman, A.; Belcourt, D.; Bennett, H.; Lazure, C.; Solomon, S.

Biochem. Biophys. Res. Commun. 173, 1161-1168, 1990

A:Title: Granulins, a novel class of peptide from leukocytes.

A:Reference number: A36698; MUID:91097544; PMID:2268320

A:Accession: A36698

A:Molecule type: protein

A:Residues: 281-336 <BAT>

A:Note: this protein was purified and characterized as granulin A

A:Accession: B36698

A:Molecule type: protein

A:Residues: 206-218, 'H', 220-233 <BA2>

A:Note: this protein was purified and characterized as granulin B

A:Accession: C36698

A:Molecule type: protein

A:Residues: 364-367, 'X', 369-385, 'H', 387-396 <BA3>

A:Note: this protein was purified and characterized as granulin C

A:Accession: D36698

A:Molecule type: protein

A:Residues: 442-446, 'XDTSS', 456-458, 'DG' <BA4>

LSU ribosomal prot  
hypothetical prote  
hypothetical prote  
porphobilinogen de  
protein R09E10.1 [  
hypothetical prote  
HYA22 protein - hu  
r40q2 protein - ri  
sucrose-phosphate  
DNA-binding protei  
related to suppres  
DNA-directed RNA p  
probable vitellogen  
ALR protein - huma  
ALR protein - huma  
bcl-2-associated p

30 36 48.6 201 2 AH3346  
31 36 48.6 302 2 T24091  
32 36 48.6 307 2 F71294  
33 36 48.6 322 2 G87257  
34 36 48.6 326 2 H88808  
35 36 48.6 336 2 T15837  
36 36 48.6 338 2 JG5707  
37 36 48.6 343 2 T03960  
38 36 48.6 348 2 S72650  
39 36 48.6 402 2 S61413  
40 36 48.6 1174 2 T49868  
41 36 48.6 1210 2 S35548  
42 36 48.6 1984 2 T13171  
43 36 48.6 4957 2 T03455  
44 36 48.6 5262 2 T03454  
45 35 47.3 41 2 C47538

A;Accession: A46705  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 18-19,'X',21-25,'X',27-29,'XX',32,'XXX',119-127,152-154,'DXK',158-161,'  
C;Superfamily: granulin

Query Match 71.6%; Score 53; DB 2; Length 589;  
Best Local Similarity 75.0%; Pred. No. 0.12;  
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RGTKCLRRAPR 14  
|||||::||  
Db 564 RGTKCLRRKIPR 575

RESULT 3  
B38128  
epithelin/granulin precursor - rat.  
C;Species: Rattus norvegicus (Norway rat)  
C;date: 10-Jul-1992;#sequence\_revision 10-Jul-1992 #text\_change 20-Aug-1999  
C;Accession: B38128; A36199; B36199; E36698; F53272  
R;Plozman, G.D.; Green, J.M.; Neubauer, M.G.; Buckley, S.D.; McDonald, V.L.; Todaro  
J. Biol. Chem. 267, 13073-13078, 1992  
A;Title: The epithelin precursor encodes two proteins with opposing activities on  
A;Reference number: A38128; MUID:92317004; PMID:1618805  
A;Accession: B38128  
A;Molecule type: mRNA  
A;Residues: 1-589 <PIG>

R;Shoyab, M.; McDonald, V.L.; Byles, C.; Todaro, G.J.; Plozman, G.D.  
Proc. Natl. Acad. Sci. U.S.A. 87, 7912-7916, 1990  
A;Title: Epithelins 1 and 2: isolation and characterization of two cysteine-rich gr  
A;Reference number: A36199; MUID:91045907; PMID:2236009  
A;Accession: A36199  
A;Molecule type: protein  
A;Residues: 280-300 <SHO>

A;Molecule type: protein  
A;Residues: 205-226 <SH2>  
R;Bateman, A.; Belcourt, D.; Bennett, H.; Lazure, C.; Solomon, S.  
Biochem. Biophys. Res. Commun. 173, 1161-1168, 1990  
A;Title: Granulins, a novel class of peptide from leukocytes.  
A;Reference number: A36698; MUID:91097544; PMID:2268320  
A;Accession: E36698  
A;Molecule type: protein  
A;Residues: 279-307,'SB',310-324,'T',326,'X',328,'Q' <BAT>  
R;Bhandari, V.; Giald, A.; Bateman, A.  
Endocrinology 133, 2682-2689, 1993  
A;Title: The complementary deoxyribonucleic acid sequence, tissue distribution, and  
A;Reference number: I53272; MUID:94062640; PMID:8243292  
A;Accession: I53272  
A;Status: preliminary; translated from GB/EMBL/DBDJ  
A;Molecule type: mRNA  
A;Residues: 1-200,'S',203-388,'M',390-589 <RES>  
A;Cross-references: GB:M97750; NID:g204223; PIDN:AAAL6903.1; PID:g204224  
C;Superfamily: granulin

Query Match 68.9%; Score 51; DB 2; Length 589;  
Best Local Similarity 66.7%; Pred.No. 0.29;  
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 RGTKCLRRAPR 14  
:|||||: ||  
Db. 564 KGTKCLRKTKPR 575

RESULT 4  
A75577  
nicotinate-nucleotide-dimethylbenzimidazole phosphoribosyltransferase - Deinococcus  
C;Species: Deinococcus radiodurans  
C;Date: 03-Dec-1999 #sequence\_revision  
C;Accession: A75577  
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, J.

RESULT 4  
A7577  
nicotinate-nucleotide-dimethylbenzimidazole phosphoribosyltransferase - *Deinococcus*  
C;Species: *Deinococcus radiodurans*  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 02-Mar-2001  
C;Accession: A7577  
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.

M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896; PMID:10567266  
A:Accession: A75577  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-365 <WHI>  
A:Cross-references: GB:AE001863; GB:AE001825; NID:96460670; PIDN:AAF12397.1; PID:9646069  
A:Experimental source: strain R1  
C:Genetics:  
A:Map position: 2  
C:Superfamily: Pseudomonas nicotinate-nucleotide-dimethylbenzimidazole phosphoribosyltra

Query Match 58.1%; Score 43; DB 2; Length 365;  
Best Local Similarity 81.8%; Pred. No. 5.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RRGTKCLRREA 12  
||||| |||||  
Db 146 RRGTRNLREA 156

RESULT 5  
T49514  
hypothetical protein B14D6.660 [imported] - Neurospora crassa  
C:Species: Neurospora crassa  
C:Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000  
A:Accession: T49514  
R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,  
submitted to the Protein Sequence Database, May 2000  
A:Reference number: ZJ5022  
A:Accession: T49514  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-327 <SCH>  
A:Cross-references: EMBL:AL356173; GSPDB:GN00116; NCSP:B14D6.660  
A:Experimental source: BAC clone B14D6; strain OR74A  
C:Genetics:  
A:Gene: NCSP:B14D6.660  
A:Map position: 6

Query Match 55.4%; Score 41; DB 2; Length 327;  
Best Local Similarity 58.3%; Pred. No. 11;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 RRGTKCLRREAP 13  
:||| |||||  
Db 275 QRGPOCLRRPP 286

RESULT 6  
T49513  
gastric mucin related protein [imported] - Neurospora crassa  
A:Alternate names: protein B14D6.650  
C:Species: Neurospora crassa  
C:Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000  
A:Accession: T49513  
R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,  
submitted to the Protein Sequence Database, May 2000  
A:Reference number: ZJ5022  
A:Accession: T49513  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-422 <SCH>  
A:Cross-references: EMBL:AL356173; GSPDB:GN00116; NCSP:B14D6.650  
A:Experimental source: BAC clone B14D6; strain OR74A  
C:Genetics:  
A:Gene: NCSP:B14D6.650  
A:Map position: 6  
A:Introns: 93/2

Query Match 55.4%; Score 41; DB 2; Length 422;  
Best Local Similarity 58.3%; Pred. No. 13;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 RRGTKCLRREAP 13  
:||| |||||  
Db 370 QRGPOCLRRPP 381

RESULT 7  
T04062  
sucrose-phosphate synthase homolog F28M11.40 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 16-Jul-1999  
A:Accession: T04062  
R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Maye,  
submitted to the Protein Sequence Database, March 1999  
A:Reference number: Z15184  
A:Accession: T04062  
A:Molecule type: DNA  
A:Residues: 1-1083 <BEV>  
A:Cross-references: EMBL:AL049487  
A:Experimental source: cultivar Columbia; BAC clone F28M11  
C:Genetics:  
A:Map position: 4  
A:Introns: 86/3; 116/3; 255/3; 322/2; 482/3; 526/3; 570/3; 629/3; 647/3; 668/3;  
A:Note: F28M11.40  
C:Superfamily: sucrose-phosphate synthase; sucrose/sucrose-phosphate synthase ho  
F:230-714/Domain: sucrose/sucrose-phosphate synthase homology <SSPS>

Query Match 55.4%; Score 41; DB 2; Length 1083;  
Best Local Similarity 61.5%; Pred. No. 31;  
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 RRGTKCLRREAP 14  
||| |||||  
Db 468 RRGVCLGRYMP 480

RESULT 8  
S06187  
RNA2 polyprotein - grapevine chrome mosaic virus  
N:Contains: coat protein  
C:Species: grapevine chrome mosaic virus  
C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 08-Oct-1999  
C:Accession: S06187  
R:Brault, V.; Hibrand, L.; Candresse, T.; le Gall, O.; Dunez, J.  
Nucleic Acids Res. 17, 7809-7819, 1989  
A:Title: Nucleotide sequence and genetic organization of Hungarian grapevine chr  
A:Reference number: S06187; MUID:90016865; PMID:2798129  
A:Accession: S06187  
A:Molecule type: mRNA  
A:Residues: 1-1324 <BRA>  
A:Cross-references: EMBL:X15163; NID:959347; PIDN:CAA33255.1; PID:959348  
A:Note: part of this sequence, including the amino end of the coat protein, was  
C:Genetics:  
A:Map position: segment S2  
A:Introns: 960/2  
C:Keywords: polyprotein

Query Match 55.4%; Score 41; DB 2; Length 1324;  
Best Local Similarity 58.3%; Pred. No. 37;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 RRGTKCLRREAP 13  
||| |||||  
Db 127 RHGTRCLRHGPP 138

RESULT 9  
T19782  
hypothetical protein C36B1.11 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T19782  
 R:Lennard, N.  
 submitted to the EMBL Data Library, September 1996  
 A:Reference number: Z19176  
 A:Accession: T19782  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-327 <WIL>  
 A:Cross-references: EMBL:Z80215; PIDN:CAB02276.1; GSPDB:GN00019; CESP:C36B1.11  
 A:Experimental source: Clone C36B1  
 C:Genetics:  
 A:Gene: CESP:C36B1.11  
 A:Map position: 1  
 A:Introns: 83/3; 125/3; 174/3; 197/1; 261/3; 323/2

Query Match 54.1%; Score 40; DB 2; Length 327;  
 Best Local Similarity 53.8%; Pred. No. 16;  
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 ARRGTKLRREAP 13  
 :||| :|||:  
 Db 274 SRPFMCIRRESP 286

## RESULT 10

G70766  
 hypothetical protein Rv2084 - Mycobacterium tuberculosis (strain H37RV)  
 C:Species: Mycobacterium tuberculosis  
 C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
 C:Accession: G70766  
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 393, 537-544, 1998  
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A:Reference number: A70500; MUID:98295987; PMID:9634230  
 A:Accession: G70766  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-378 <COL>  
 A:Cross-references: GB:Z73966; GB:AL123456; NID:G3261577; PIDN:CAA98196.1; PID:gl370250  
 A:Experimental source: strain H37RV  
 C:Genetics:  
 A:Gene: Rv2084  
 C:Superfamily: Mycobacterium tuberculosis hypothetical protein Rv2084

Query Match 54.1%; Score 40; DB 2; Length 378;  
 Best Local Similarity 71.4%; Pred. No. 18;  
 Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ARRGTKLRREAP 14  
 :||||| :|||||  
 Db 306 ARRGTRGRRCAPR 319

## RESULT 11

T50868  
 exopolysphatase [imported] - Rubrivivax gelatinosus  
 C:Species: Rubrivivax gelatinosus  
 C:Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 18-Aug-2000  
 C:Accession: T50868  
 R:Nagashima, K.V.; Igarashi, N.; Harada, J.; Nagashima, S.; Matsuura, K.; Shimada, K.  
 submitted to the EMBL Data Library, November 1999  
 A:Description: Determination of Nucleotide Sequences of Rubrivivax gelatinosus Photosynt  
 A:Reference number: Z25270  
 A:Accession: T50868  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-500 <NAG>  
 A:Cross-references: EMBL:AB034704; PIDN:BAA94021.1

A:Experimental source: strain ILL44  
 C:Genetics:  
 A:Gene: ppx  
 C:Superfamily: exopolysphatase

Query Match 54.1%; Score 40; DB 2; Length 500;  
 Best Local Similarity 61.5%; Pred. No. 24;  
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 RRGTKLRREAPR 14  
 :||| :|||:  
 Db 61 QRGLECLRREFAAR 73

## RESULT 12

H84368  
 MCM / cell division control protein 21 [imported] - Halobacterium sp. NRC-1  
 C:Species: Halobacterium sp. NRC-1  
 C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
 C:Accession: H84368  
 R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; La;  
 ; Leithausen, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.;  
 Jung, K.H.; Alam, M.; Freitas, T.  
 Proc Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
 A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebbardt, H.; Lowe, T.  
 A:Title: Genome sequence of Halobacterium species NRC-1.  
 A:Reference number: A84160; MUID:20504483; PMID:11016950  
 A:Accession: H84368  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-831 <STO>  
 A:Cross-references: GB:AE004437; NID:gl0581599; PIDN:AAG20316.1; GSPDB:GN00138  
 C:Genetics:  
 A:Gene: mcm

Query Match 54.1%; Score 40; DB 2; Length 831;  
 Best Local Similarity 66.7%; Pred. No. 37;  
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 RRGTKLRREAP 13  
 :||| :|||:  
 Db 338 RRATKVKREAP 349

## RESULT 13

T08723  
 hypothetical protein DKFZp566D193.1 - human (fragment)  
 C:Species: Homo sapiens (man)  
 C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 13-Aug-1999  
 C:Accession: T08723  
 R:Koehrer, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
 submitted to the Protein Sequence Database, March 1999  
 A:Reference number: Z16473  
 A:Accession: T08723  
 A:Molecule type: mRNA  
 A:Residues: 1-118 <KOE>  
 A:Cross-references: EMBL:AL050051  
 A:Experimental source: fetal kidney; clone DKFZp566D193  
 C:Genetics:  
 A:Note: DKFZp566D193.1

Query Match 52.7%; Score 39; DB 2; Length 118;  
 Best Local Similarity 77.8%; Pred. No. 9.9;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 GTKCLREA 12  
 :||| :|||:  
 Db 79 GTSLRQEA 87

## RESULT 14

D70775  
 hypothetical protein Rv1312 - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis  
 C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
 C:Accession: D70775  
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Cole, S.T.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998  
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A:Reference number: A70500; PMID:98295987; PMID:9634230  
 A:Accession: D70775  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-147 <COL>  
 A:Cross-references: GB:273419; GB:AL123456; NID:93261573; PIDN:CAA97745.1; PID:gl322436  
 A:Experimental source: strain H37Rv  
 C:Genetics:  
 A:Gene: Rv1312

Query Match 52.7%; Score 39; DB 2; Length 147;  
 Best Local Similarity 50.0%; Pred. No. 12;  
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 ARGTCCLRRAPR 14  
 :|||:|:|  
 Db 78 SRGVEIISRRAPR 91

## RESULT 15

C75558

acetyl-CoA carboxylase, biotin carboxyl carrier protein - Deinococcus radiodurans (strain C75558)  
 C:Species: Deinococcus radiodurans  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000  
 C:Accession: C75558  
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999  
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
 A:Reference number: A75250; PMID:20036896; PMID:10567266  
 A:Accession: C75558  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-187 <WHI>  
 A:Cross-references: GB:AE001874; GB:AE000513; NID:96457778; PIDN:AAF09708.1; PID:9645778  
 A:Experimental source: strain R1  
 C:Genetics:  
 A:Gene: DR0118  
 A:Map position: 1  
 C:Superfamily: biotin carboxyl carrier protein; lipoyl/biotin-binding homology

Query Match 51.4%; Score 38; DB 2; Length 187;  
 Best Local Similarity 53.8%; Pred. No. 22;  
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 RRGTCCLRRAPR 14  
 ||:|:|  
 Db 37 RRPQCLPRSAQ 49

Search completed: July 8, 2003, 16:29:31  
 Job time : 16.1795 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using: sw model

Run on: July 8, 2003, 16:16:30 ; Search time 6.82051 Seconds  
(without alignments)  
85.136 Million cell updates/sec

Title: US-09-824-647-7

Perfect score: 74

Sequence: 1 ARRTKCLRREAPR 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68	91.9	593	1 GRN_HUMAN	P28799 h granulin
2	53	71.6	589	1 GRN_MOUSE	P28798 mus musculus
3	51	68.9	588	1 GRN_RAT	P23785 r. granulin
4	43	58.1	365	1 COBT_DEIRA	Q91978 deinococcus
5	41	55.4	1324	1 POL2_GCMV	P13026 hungaricus g
6	40	54.1	380	1 YR84_MYCTU	Q10692 mycobacteri
7	39	52.7	147	1 YD12_MYCTU	Q10620 mycobacteri
8	39	52.7	1121	1 DPOL_ADEG1	Q64751 avian adeno
9	38	51.4	118	1 YE26_RALSO	Q8xz99 ralstonia s
10	38	51.4	611	1 OM70_MOUSE	Q9czw5 mus musculus
11	38	51.4	648	1 Y084_HUMAN	Q14699 homo sapien
12	38	51.4	1068	1 SPS_MAIZE	P31927 zea mays (m
13	37	50.0	533	1 ALG6_ARATH	Q9ff17 arabidopsis
14	37	50.0	1049	1 SPS_ORYSA	Q43802 oryza sativ
15	37	50.0	1081	1 SPS2_CRAPL	Q04933 craterostig
16	36	48.6	307	1 Y678_TREPA	Q83684 treponema p
17	36	48.6	340	1 NIF1_HUMAN	Q15194 homo sapien
18	36	48.6	1210	1 RPB2_SCHPO	Q02061 schizosacch
19	36	48.6	1984	1 YL_DROME	P98163 drosophila
20	35	47.3	41	1 BAXC_HUMAN	Q07815 homo sapien
21	35	47.3	158	1 VH21_SFVKA	Q9q909 shape fibro
22	35	47.3	287	1 YKFA_ECOLI	P75678 escherichia
23	35	47.3	337	1 VINT_BP22	P36932 bacterioph
24	35	47.3	441	1 GDF9_MOUSE	Q07105 mus muscul
25	35	47.3	673	1 Y552_HUMAN	Q60299 homo sapien
26	35	47.3	821	1 CTRL_ARATH	Q05609 arabidopsis
27	35	47.3	843	1 BLVR_BOVIN	Q03368 bos taurus
28	35	47.3	916	1 CHSB_EMENI	Q00757 emericella
29	35	47.3	2411	1 DAB_DROME	P98081 drosophila
30	34	45.9	67	1 BD03_HUMAN	P81534 homo sapien
31	34	45.9	132	1 Y4LK_RHISN	P55551 rhizobium s
32	34	45.9	147	1 YD12_MYCLE	P53432 mycobacteri
33	34	45.9	182	1 NQ09_THETH	Q56224 thermus the

34 34 45.9 289 1 HEM3\_METTH  
35 34 45.9 410 1 ICP0\_PRVIF  
36 34 45.9 506 1 PPX\_PSEAE  
37 34 45.9 841 1 NEK4\_HUMAN  
38 34 45.9 1113 1 MGA2\_YEAST  
39 34 45.9 1171 1 NTFJ\_KLEPN  
40 34 45.9 1239 1 V120\_EBV  
41 34 45.9 1458 1 PHLX\_RABIT  
42 34 45.9 2212 1 RRLP\_EBOZM  
43 34 45.9 2254 1 CCAG\_RAT  
44 34 45.9 2377 1 CCAG\_HUMAN  
45 34 45.9 2704 1 G168\_PARPR

O26960 methanobact  
P29129 pseudorabie  
Q9zn70 pseudomonas  
P51957 homo sapien  
P40578 saccharomyc  
P03833 klebsiella  
P03189 epstein-bar  
Q05087 oryctolagus  
Q05318 ebola virus  
O54898 rattus norv  
O43497 homo sapien  
P17053 paramecium

#### ALIGNMENTS

#### RESULT 1

GRN\_HUMAN STANDARD; PRT: 593 AA.  
AC P28799; P23781; P23782; P23783; P23784; Q9BWE7;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Granulins precursor (Acrogranin) [Contains: Paragranulin; Granulin 1  
(Granulin G); Granulin 2 (Granulin F); Granulin 3 (Granulin B);  
Granulin 4 (Granulin A); Granulin 5 (Granulin C); Granulin 6 (Granulin  
D); Granulin 7 (Granulin E)].  
GN GRN.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC TISSUE=Bone marrow;  
RX MEDLINE=92179253; PubMed=1542665;  
RA Bhandari V., Falfree R.G.E., Bateman A.;  
RT "Isolation and sequence of the granulin precursor cDNA from human  
bone marrow reveals tandem cysteine-rich granulin domains.";  
RL Proc. Natl. Acad. Sci. U.S.A. 89:1715-1719(1992).  
RN [2]  
RX REVISIONS, SEQUENCE FROM N.A.  
RA MEDLINE=93038704; PubMed=1417868;  
RA Bhandari V., Bateman A.;  
RT "Structure and chromosomal location of the human granulin gene.";  
RL Biochem. Biophys. Res. Commun. 188:57-63(1992).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RX MEDLINE=92317004; PubMed=1618805;  
RA Plozman G.D., Green I.M., Neubauer M.G., Buckley S.D., McDonald V.L.,  
RTodaro G.J., Shoyab M.;  
RT "The epithelin precursor encodes two proteins with opposing activities  
on epithelial cell growth.";  
RL J. Biol. Chem. 267:13073-13078(1992).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Yu W., Gibbs R.A.;  
RT Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
RC TISSUE=Cervix, and Lung;  
RA Strausberg R.;  
RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE OF 206-233; 281-336; 364-396 AND 442-447.  
RC TISSUE=Leukocyte;  
RA Bateman A., Belcourt D.R., Bennett H.P., Lazure C., Solomon S.;  
RT "Granulins, a novel class of peptide from leukocytes.";  
RL Biochem. Biophys. Res. Commun. 173:1161-1168(1990).

CC -!- FUNCTION: GRANULINS HAVE POSSIBLE CYTOKINE-LIKE ACTIVITY. THEY MAY  
CC PLAY A ROLE IN INFLAMMATION, WOUND REPAIR, AND TISSUE REMODELING.  
CC -!- FUNCTION: GRANULIN A PROMOTES PROLIFERATION OF THE EPITHELIAL  
CC CELL LINE A431 IN CULTURE WHILE GRANULIN B ACTS AS AN ANTAGONIST  
CC TO GRANULIN A, INHIBITING THE GROWTH.  
CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; may be  
CC produced by alternative splicing.  
CC -!- TISSUE SPECIFICITY: IN MYELOGENOUS LEUKEMIC CELL LINES OF  
CC PROMONOCYTIC, PROMYELOCYTIC, AND PROERYTHROID LINEAGE, IN  
CC FIBROBLASTS, AND VERY STRONGLY IN EPITHELIAL CELL LINES. PRESENT  
CC IN INFLAMMATORY CELLS AND BONE MARROW. HIGHEST LEVELS IN KIDNEY.  
CC -!- PTM: GRANULINS ARE DISULFIDE BRIDGED.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC

DR EMBL; W75161; AAA58617.1; ALT\_SEQ.  
DR EMBL; X62320; CAA44196.1; -  
DR EMBL; AF055008; AAC09359.1; -  
DR EMBL; BC000324; AAH00324.1; -  
DR EMBL; BC010577; AAH10577.1; -  
DR PIR; A38118; GYHU.  
DR PIR; D36698; D36698.  
DR PIR; JC1284; JC1284.  
DR GENE; HGNC:4601; GRN.  
DR MIM; 138945; -  
DR InterPro: IPR000118; Granulin.  
DR Pfam: PF00396; granulin; 7.  
DR SMART; SM00277; GRAN; 7.  
DR PROSITE; PS00799; GRANULINS; 7.  
KW Cytokine; Repeat; Glycoprotein; Signal; Alternative splicing;  
KW Polymorphism.  
FT SIGNAL 1 17 POTENTIAL.  
FT CHAIN 18 593 ACROGRANIN.  
FT PEPTIDE 18 747 PARAGRANULIN.  
FT PEPTIDE 758 2113 GRANULIN 1.  
FT PEPTIDE 7123 2179 GRANULIN 2.  
FT PEPTIDE 206 261 GRANULIN 3.  
FT PEPTIDE 281 336 GRANULIN 4.  
FT PEPTIDE 364 747 GRANULIN 5.  
FT PEPTIDE 442 7496 GRANULIN 6.  
FT PEPTIDE 7518 7573 GRANULIN 7.  
FT CARBOHYD 118 118 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 236 236 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 265 265 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 368 368 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 530 530 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT VARSPIC 377 531 MISSING (IN ISOFORM 2).  
FT VARIANT 454 454 G -> Q.  
FT FTID-VAR 003445.  
FT CONFLICT 219 219 S -> H (IN REF. 6).  
FT CONFLICT 386 386 W -> H (IN REF. 6).  
SQ SEQUENCE 593 AA; 63473 MW; 4E402BDB16DE2819 CRC64;

Query Match 91.9%; Score 68; DB 1; Length 593;  
Best Local Similarity 92.9%; Pred. No. 9.8e-05;  
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ARGTKCLRRREAPR 14  
Db 566 AARGTKCLRRREAPR 579

RESULT 2  
GRN\_MOUSE STANDARD; PRT; 589 AA.  
AC P28798;  
DT 01-DEC-1992 (Rel. 24, Created)

DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Granulins precursor (Acrogranin) [Contains: Granulin 1; Granulin 2;  
DE Granulin 3; Granulin 4; Granulin 5; Granulin 6; Granulin 7].  
GN GRN.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_Taxid:10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-93245991; PubMed-8482392;  
RA Baba T., Nemoto H., Watanabe K., Arai Y., Gerton G.L.;  
RT "Exon/Intron organization of the gene encoding the mouse  
RT epithelin/granulin precursor (acrogranin).";  
RL FEBS Lett. 322:89-94(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-92317004; PubMed-1618805;  
RA Plowman G.D., Green I.M., Neubauer M.G., Buckley S., McDonald V.L.,  
RA Todaro G.I., Shoyab M.;  
RT "The epithelin precursor encodes two proteins with opposing  
RT activities on epithelial cell growth.";  
RL J. Biol. Chem. 267:13073-13078(1992).  
CC -!- FUNCTION: GRANULINS HAVE POSSIBLE CYTOKINE-LIKE ACTIVITY. THEY MAY  
CC PLAY A ROLE IN INFLAMMATION, WOUND REPAIR, AND TISSUE REMODELING.  
CC -!- TISSUE SPECIFICITY: UBIQUITOUS.  
CC -!- PTM: GRANULINS ARE DISULFIDE BRIDGED.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC

DR EMBL; D16195; BAA03736.1; -  
DR EMBL; M86736; AAA37191.1; -  
DR EMBL; X62321; CAA44197.1; -  
DR MGD; MGI:95832; Grn.  
DR InterPro: IPR000118; Granulin.  
DR Pfam: PF00396; granulin; 7.  
DR SMART; SM00277; GRAN; 7.  
DR PROSITE; PS00799; GRANULINS; 7.  
KW Cytokine; Repeat; Glycoprotein; Signal.  
FT SIGNAL 1 17 POTENTIAL.  
FT CHAIN 18 589 ACROGRANIN.  
FT PEPTIDE 18 589 ACROGRANIN.  
FT PEPTIDE 258 2113 GRANULIN 1.  
FT PEPTIDE 7122 2178 GRANULIN 2.  
FT PEPTIDE 205 260 GRANULIN 3.  
FT PEPTIDE 280 334 GRANULIN 4.  
FT PEPTIDE 362 7414 GRANULIN 5.  
FT PEPTIDE 440 7493 GRANULIN 6.  
FT PEPTIDE 7517 7568 GRANULIN 7.  
FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 373 373 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 526 526 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CONFLICT 350 350 L -> R (IN REF. 2).  
SQ SEQUENCE 589 AA; 63458 MW; 1DE8229C413CB787 CRC64;

Query Match 71.6%; Score 53; DB 1; Length 589;  
Best Local Similarity 75.0%; Pred. No. 0.049;  
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 RGTKCLRRREAPR 14  
Db 564 RGTKCLRRREAPR 575

RESULT 3

GRN\_RAT ID GRN\_RAT STANDARD; PRT; 588 AA.

AC P23785;

DT 01-NOV-1991 (Rel. 20, Created)

DT 01-DEC-1992 (Rel. 24, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Granulins precursor (Acrogranin) (Contains: Granulin 1 (Granulin G); Granulin 2 (Granulin F); Granulin 3 (Granulin B) (Epithelin 2); Granulin 4 (Granulin A) (Epithelin 1); Granulin 5 (Granulin C); Granulin 6 (Granulin D); Granulin 7 (Granulin E)).

GN GRN.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.

OX NCBI\_Taxid=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Kidney;

RX MEDLINE=94062640; PubMed=8243292;

RA Bhandari V., Giald A., Bateman A.;

RT "The complementary deoxyribonucleic acid sequence, tissue distribution, and cellular localization of the rat granulins precursor."

RT precursor."

RL Endocrinology 133:2682-2689(1993).

RN [2]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 204-259 AND 278-334.

RC TISSUE=Kidney;

RX MEDLINE=92317004; PubMed=1618805;

RA Plovman G.D., Green J.M., Neubauer M.G., Buckley S.D., McDonald V.L., Todaro G.J., Shoyab M.;

RT "The epithelin precursor encodes two proteins with opposing activities on epithelial cell growth."

RT activities on epithelial cell growth."

RL J. Biol. Chem. 267:13073-13078(1992).

RN [3]

RP SEQUENCE OF 204-225 AND 279-299.

RX MEDLINE=91045907; PubMed=2236009;

RA Shoyab M., McDonald V.L., Byles C., Todaro G.J., Plovman G.D.;

RT "Epithelins 1 and 2: isolation and characterization of two cysteine-rich growth-modulating proteins."

RT cysteine-rich growth-modulating proteins."

RL Proc. Natl. Acad. Sci. U.S.A. 87:7912-7916(1990).

RN [4]

RP SEQUENCE OF 278-328.

RC TISSUE=Bone marrow;

RX MEDLINE=91097544; PubMed=2268320;

RA Bateman A., Belcourt D.R., Bennett H.P., Lazure C., Solomon S.;

RT "Granulins, a novel class of peptide from leukocytes."

RL Biochem. Biophys. Res. Commun. 173:1161-1168(1990).

CC -1- FUNCTION: GRANULINS HAVE POSSIBLE CYTOKINE-LIKE ACTIVITY. THEY MAY PLAY A ROLE IN INFLAMMATION, WOUND REPAIR, AND TISSUE REMODELING.

CC -1- TISSUE SPECIFICITY: UBIQUITOUS; MOST ABUNDANT IN THE SPLEEN AND SEVERAL TISSUES OF ENDOCRINE SIGNIFICANCE.

CC -1- PTM: GRANULINS ARE DISULFIDE BRIDGED.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; M97750; AA16903.1; -

DR EMBL; X62322; CA44198.1; -

DR PIR; A36199; A36199.

DR PIR; B36199; B36199.

DR PIR; E36698; E36698.

DR InterPro; IPR000118; Granulin.

DR Pfam; PF00396; granulin; 7.

DR SMART; SM00277; GRAN; 7.

DR PROSITE; PS00799; GRANULINS; 7.

DR Cytokine; Repeat; Glycoprotein; Signal.

FT SIGNAL 1 17 POTENTIAL..

FT CHAIN 18 588 ACROGRANIN.

FT PEPTIDE 58 113 GRANULIN 1.

FT PEPTIDE 122 178 GRANULIN 2.

FT PEPTIDE 204 259 GRANULIN 3.

FT PEPTIDE 278 334 GRANULIN 4.

FT PEPTIDE 361 413 GRANULIN 5.

FT PEPTIDE 438 492 GRANULIN 6.

FT PEPTIDE 512 567 GRANULIN 7.

FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 372 372 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 525 525 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CONFLICT 201 201 S -> FP (IN REF. 2).

FT CONFLICT 307 308 TK -> SB (IN REF. 4).

FT CONFLICT 324 324 Q -> T (IN REF. 4).

FT CONFLICT 388 388 M -> I (IN REF. 2).

SQ SEQUENCE 588 AA; 63369 MW; 113D434F7E099831 CRC64;

Query Match 68.9%; Score 51; Length 588;

Best Local Similarity 66.7%; Pred. No. 0.11;

Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 3 RGTGCLREAPR 14

Db 563 RGTGCLREKTPR 574

RESULT 4

COBT\_DEIRA

ID COBT\_DEIRA STANDARD; PRT; 365 AA.

AC Q9RYR8;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Nicotinate-nucleotide--dimethylbenzimidazole phosphoribosyltransferase (EC 2.4.2.21) (NN:DBI PRT) (N(1)-alpha-phosphoribosyltransferase).

DE COBT OR DRA0240.

GN Deinococcus radiodurans.

OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;

OC Deinococcaceae; Deinococcus.

OX NCBI\_Taxid=1299;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=RL;

RX MEDLINE=20036896; PubMed=10567266;

RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D., Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L., Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M., Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C., Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D., Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C., Fraser C.M.;

RA "Genome sequence of the radioresistant bacterium Deinococcus radiodurans RL."

RT Science 286:1571-1577(1999).

CC -1- FUNCTION: CATALYZES THE SYNTHESIS OF ALPHA-RIBAZOLE-5'-PHOSPHATE FROM NICOTINATE MONONUCLEOTIDE (NAMN) AND 5,6-DIMETHYLBENZIMIDAZOLE (DMB).

CC -1- CATALYTIC ACTIVITY: Beta-nicotinate D-ribonucleotide + dimethylbenzimidazole = nicotinate + N(1)-(5-phospho-alpha-D-ribose)l-5,6-dimethylbenzimidazole.

CC -1- PATHWAY: Cobalamin biosynthesis.

CC -1- SIMILARITY: BELONGS TO THE COBT FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AE001863; AAF12397.1; -

DR HSSP; Q05603; 1D0V.

DR TIGR; DRA0240; -



```
DR InterPro: IPR003200; NN:DBI_PRT.
DR Pfam: PF02277; DBI_PRT: 1.
KW Cobalamin biosynthesis; Porphyrin biosynthesis; Transferase;
KW Glycosyltransferase; Complete proteome.
FT ACT_SITE 330 BASE (BY SIMILARITY).
SQ SEQUENCE 365 AA; 37265 MW; 81A37D3194F2EC7D CRC64;

Query Match 58.1%; Score 43; DB 1; Length 365;
Best Local Similarity 81.8%; Pred. No. 1.9;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRGTRCLRREA 12
DB 146 RRGTRNLRREA 156
||||| |||||

RESULT 5
POL2_GCMV STANDARD; PRT; 1324 AA.
AC P13026;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE RNA2 polyprotein (147 kDa protein) (Contains: Coat protein; 84 kDa
DE protein).
OS Hungarian grapevine chrome mosaic virus (GCMV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Comoviridae;
OC Nepovirus.
OX NCBI_Taxid:12273;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 814-822.
RX MEDLINE-90016865; PubMed-2798129;
RA Braut V., Hibrand L., Candresse T., le Gall O., Dunez J.;
RT "Nucleotide sequence and genetic organization of Hungarian grapevine
RT chrome mosaic nepovirus RNA2.";
RL Nucleic Acids Res. 17:7809-7823(1989).
CC -!- FUNCTION: THE PROTEIN LOCATED AT THE N-TERMINAL END OF THE
CC NEPOVIRUS POLYPROTEIN COULD BE REQUIRED FOR NEMATODE TRANSMISSION
CC OF THE VIRUS.
CC -!- PTM: THE RNA2 POLYPROTEIN IS CLEAVED BY AN RNA1-ENCODED PROTEASE
CC TO YIELD THE MATURE COAT PROTEIN AND AN 84 KDA PROTEIN WHICH IS
CC FURTHER CLEAVED INTO TWO PRODUCTS OF APPROXIMATELY 46 AND 48 KDA.
CC -!- SIMILARITY: TO THE RNA2 POLYPROTEIN OF OTHER NEPOVIRUSES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR ENBL: X15163; CAA33255.1;
DR PIR: S06187;
DR InterPro: IPR005054; Nepo_coat.
DR InterPro: IPR005305; Nepo_coat_C.
DR InterPro: IPR005306; Nepo_coat_N.
DR Pfam: PF03391; Nepo_coat: 1.
DR Pfam: PF03688; Nepo_coat_C: 1.
DR Pfam: PF03689; Nepo_coat_N: 1.
DR Polyprotein; Coat protein.
KW MOD_RES 811 811 BLOCKED (PROBABLE).
FT DOMAIN 1 810 84 KDA PROTEIN.
FT DOMAIN 811 1324 COAT PROTEIN.
FT VARIANT 400 400 N -> D.
FT VARIANT 1006 1006 I -> T.
SQ SEQUENCE 1324 AA; 147804 MW; 302C286B2EA1756A CRC64;

Query Match 55.4%; Score 41; DB 1; Length 1324;
Best Local Similarity 58.3%; Pred. No. 16;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 RRGTRCLRREAP 13
```

```
DB 127 RHGTRCLRGGP 138
|||||
RESULT 6
YR84_MYCTU STANDARD; PRT; 380 AA.
AC Q10692;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2001 (Rel. 41, Last annotation update)
DE Hypothetical protein RV2084.
GN RV2084 OR MT2146 OR MTCY49.23.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_Taxid:1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H37Rv;
RX MEDLINE-98295987; PubMed-9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RL complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR ENBL: Z73966; CAA98196.1; ALT_INIT.
DR EMBL: AE007064; AAK46428.1;
DR TIGR: MT2146;
DR Tuberculist; RV2084;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 380 AA; 41314 MW; D61A51A1BC8775EE CRC64;

Query Match 54.1%; Score 40; DB 1; Length 380;
Best Local Similarity 71.4%; Pred. No. 6.8;
Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ARRGTKCLRREAP 14
||||| |||||
DB 308 ARRGTSRRRCAPR 321

RESULT 7
YD12_MYCTU STANDARD; PRT; 147 AA.
ID YD12_MYCTU
AC Q10620;
DT 01-OCT-1996 (Rel. 34, Created)
```



CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL; AL646064; CAD15128.1;  
 DR InterPro; IPR005346; UPF0125;  
 DR Pfam; PF03658; UPF0125; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 118 AA; 12743 MW; C0594E56E184E769 CRC64;

Query Match 51.4%; Score 38; DB 1; Length 118;  
 Best Local Similarity 66.7%; Pred. No. 4.8;  
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 2 RRGTKCLRRAP 13  
 Db 100 REGLKWLNRAP 111

RESULT 10  
 OM70\_MOUSE  
 ID OM70\_MOUSE STANDARD; PRT; 611 AA.  
 AC Q9CZWS;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Mitochondrial precursor proteins import receptor (Translocase of outer membrane TOM70).  
 DE TOMM70A OR D16WSU109E.  
 GN Mus musculus (Mouse).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Embryo;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamazaki I.,  
 RA Saito T., Okazaki H., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda Y.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bona M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
 RA Hayashizaki Y.;  
 RT \*Functional annotation of a full-length mouse cDNA collection.\*;  
 RL Nature 409:685-690(2001).

CC -1- FUNCTION: RECEPTOR THAT ACCELERATES THE IMPORT OF ALL  
 CC MITOCHONDRIAL PRECURSOR PROTEINS (BY SIMILARITY).  
 CC -1- SUBUNIT: FORMS PART OF THE RECEPTOR COMPLEX THAT CONSISTS OF AT  
 CC LEAST 8 DIFFERENT PROTEINS (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial  
 CC outer membrane (By similarity).  
 CC -1- SIMILARITY: CONTAINS 10 TPR REPEATS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AK012084; BAB28018.1; -

DR HSPSP; P53041; IAL17.  
 DR MGI:106295; D16WSU109e.  
 DR InterPro; IPR001440; TPR.  
 DR Pfam; PF00515; TPR; 10.  
 DR SMART; SM00028; TPR; 10.  
 KW Mitochondrion; Outer membrane; Transmembrane; Repeat; TPR repeat.  
 FT DOMAIN 1 9  
 FT TRANSMEM 10 30  
 FT DOMAIN 31 611  
 FT REPEAT 117 150  
 FT REPEAT 156 189  
 FT REPEAT 297 330  
 FT REPEAT 332 365  
 FT REPEAT 370 403  
 FT REPEAT 404 437  
 FT REPEAT 445 478  
 FT REPEAT 479 512  
 FT REPEAT 514 547  
 FT REPEAT 548 581  
 SQ SEQUENCE 611 AA; 67521 MW; 486FB79FC4CE5B4F CRC64;

Query Match 51.4%; Score 38; DB 1; Length 611;  
 Best Local Similarity 41.7%; Pred. No. 25;  
 Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Oy 2 RRGTKCLRRAP 13  
 Db 375 KRGTMCQOQOP 386

RESULT 11  
 Y084\_HUMAN  
 ID Y084\_HUMAN STANDARD; PRT; 648 AA.  
 AC Q14699;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Hypothetical protein KIAA0084 (HA2022) (Fragment).  
 GN KIAA0084.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Bone marrow;  
 RX MEDLINE=95308325; PubMed=778527;  
 RA Nagase T., Miyajima N., Tanaka A., Suzuki T., Seki N., Sato S.,  
 RA Tabata S., Ishikawa K.-I., Kavarabayasi Y., Kotani H., Nomura N.;  
 RT \*Prediction of the coding sequences of unidentified human genes. III.  
 RT The coding sequences of 40 new genes (KIAA0081-KIAA0120) deduced by  
 RT analysis of cDNA clones from human cell line KG-1.\*;  
 RL DNA Res. 2:37-43(1995).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; D42043; BAA07644.1; -  
 CC Hypothetical protein.

FW NON\_TER 1  
 SQ SEQUENCE 648 AA; 70463 MW; 88F68A62D6C1CDAL CRC64;

Query Match 51.4%; Score 38; DB 1; Length 648;  
 Best Local Similarity 58.3%; Pred. No. 27;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Oy 1 ARGTCKCLRRAP 12  
 Db 1 ARGTCKCLRRAP 12

```

Db          46 SRDGTCLREPA 57

RESULT 12
SPS_MAIZE
ID   SPS_MAIZE      STANDARD;      PRT; 1068 AA.
AC   P31927;
DT   01-JUL-1993 (Rel. 26, Created)
DE   01-JUL-1993 (Rel. 26, Last sequence update)
DE   01-JUN-1994 (Rel. 29, Last annotation update)
DE   Sucrose-phosphate synthase (EC 2.4.1.14) (UDP-glucose-fructose-
DE   phosphate glucosyltransferase).
GN   SPS.
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC   Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 71-74; 206-212; 471-481 AND
RP   872-892.
RC   STRAIN-CV. PIONEER 3184; TISSUE-Leaf;
RC   MEDLINE=92338837; PubMed=1840396;
RA   Worrell A.C., Bruneau J.-M., Summerfelt K., Boersig M., Voelker T.A.;
RT   "Expression of a maize sucrose phosphate synthase in tomato alters
RT   leaf carbohydrate partitioning.";
RL   Plant Cell 3:1121-1130(1991).
CC   -1- FUNCTION: INVOLVED IN THE REGULATION OF CARBON PARTITIONING IN
CC   THE LEAVES OF PLANTS. MAY REGULATE THE SYNTHESIS OF SUCROSE AND
CC   THEREFORE PLAY A MAJOR ROLE AS A LIMITING FACTOR IN THE EXPORT OF
CC   PHOTOSYNTHETATES OUT OF THE LEAF.
CC   -1- CATALYTIC ACTIVITY: UDP-glucose + D-fructose 6-phosphate = UDP +
CC   sucrose 6-phosphate.
CC   -1- ENZYME REGULATION: ACTIVITY REGULATED BY PHOSPHORYLATION AND
CC   MODERATED BY CONCENTRATION OF METABOLITES AND LIGHT.
CC   -1- PATHWAY: Sucrose synthesis.
CC   -1- SUBUNIT: HOMODIMER OR HOMOTETRAMER.
CC   -1- DEVELOPMENTAL STAGE: GERMINATING SEEDS OR MATURE LEAVES.
CC   -1- PFM: PHOSPHORYLATED. HOWEVER, PHOSPHORYLATION IS NOT ESSENTIAL FOR
CC   ENZYME FUNCTION.
CC   -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 1.
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; M97550; AAA33513.1; -
DR   PIR; JQ1329; JQ1329.
DR   MaizeDB; 25294; -
DR   InterPro; IPR001296; Glycos_transf_1.
DR   Pfam; PF00534; Glycos_transf_1; 1.
KW   Transferase; Glycosyltransferase; Phosphorylation.
FT   DOMAIN 25 31 POLY-GLY.
SQ   SEQUENCE 1068 AA; 118575 MW; 074679B5E9A1D282 CRC64;

Query Match          51.4%; Score 38; DB 1; Length 1068;
Best Local Similarity 57.1%; Pred. No. 44;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY   1-ARRGTGKCLRREAP 14
      |||||
Db   419 ARRGVSGHGYWPR 432

RESULT 13
ALG6_ARATH
ID   ALG6_ARATH      STANDARD;      PRT; 533 AA.
AC   Q9FF17;
DT   15-JUN-2002 (Rel. 41, Created)

```

```

DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Probable dolichyl pyrophosphate Man9GlcNAc2 alpha-1,3-
DE   glucosyltransferase (EC 2.4.1.-) (Dolichyl-P-Glc:Man9GlcNAc2-pp-
DE   dolichyl glucosyltransferase).
GN   AT5G38460 OR MX110.19.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RP   STRAIN-CV. Columbia;
RC   MEDLINE=97471969; PubMed=9330910;
RA   Sato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M.,
RA   Miyajima N., Tabata S.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence
RT   features of the 1.6 Mb regions covered by twenty physically assigned
RT   P1 clones.";
RL   DNA Res. 4:215-230(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RP   STRAIN-CV. Columbia;
RA   Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT   "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
RT   SSP consortium (Salk/Stanford/PGECC).";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC   -1- FUNCTION: Adds the first glucose residue to the lipid-linked
CC   oligosaccharide precursor for N-linked glycosylation. Transfers
CC   glucose from dolichyl phosphate glucose (Dol-P-Glc) onto the
CC   lipid-linked oligosaccharide Man(9)GlcNAc(2)-pp-Dol (By
CC   similarity).
CC   -1- PATHWAY: Glycosylation.
CC   -1- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC   reticulum (Potential).
CC   -1- SIMILARITY: BELONGS TO THE ALG6/ALG8 GLYCOSYLTRANSFERASE FAMILY.
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; AB005248; BAB09358.1; -
DR   EMBL; AY035066; AAK59571.1; -
DR   EMBL; AY056364; AAL07250.1; -
DR   InterPro; IPR004856; Alg6_Alg8.
DR   Pfam; PF03155; Alg6_Alg8; 1.
KW   Hypothetical protein; Transferase; Glycosyltransferase; Transmembrane;
KW   Endoplasmic reticulum.
FT   TRANSMEM 42 64 POTENTIAL.
FT   TRANSMEM 149 169 POTENTIAL.
FT   TRANSMEM 184 204 POTENTIAL.
FT   TRANSMEM 214 234 POTENTIAL.
FT   TRANSMEM 264 284 POTENTIAL.
FT   TRANSMEM 360 380 POTENTIAL.
FT   TRANSMEM 422 442 POTENTIAL.
FT   TRANSMEM 463 483 POTENTIAL.
FT   TRANSMEM 491 511 POTENTIAL.
SQ   SEQUENCE 533 AA; 61112 MW; 79DEC123A6387F26 CRC64;

Query Match          50.0%; Score 37; DB 1; Length 533;
Best Local Similarity 75.0%; Pred. No. 33;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY   6 KCLRREAP 13
      |||||
Db   256 KCLRREKSP 263

```

## RESULT 14

SPS\_ORVSA STANDARD; PRT; 1049 AA.

AC Q43802;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Sucrose-phosphate synthase (EC 2.4.1.14)

DE (UDP-glucose-fructose-phosphate glucosyltransferase):

OS Oryza sativa (Rice).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzeae; Oryza.

OX NCBI\_TaxID=4530;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Indica-IR36; TISSUE=Leaf;

RX MEDLINE=96235138; PubMed=8666248;

RA Valdez-Alarcon J.J., Ferrando M., Jimenez-Morales B.,

RA Herrera-Estrella L.;

RT "Characterization of a rice sucrose-phosphate synthase-encoding gene.";

RL Gene 170:217-222(1996).

CC -!- FUNCTION: INVOLVED IN THE REGULATION OF CARBON PARTITIONING IN THE LEAVES OF PLANTS. MAY REGULATE THE SYNTHESIS OF SUCROSE AND THEREFORE PLAY A MAJOR ROLE AS A LIMITING FACTOR IN THE EXPORT OF PHOTOASSIMILATES OUT OF THE LEAF.

CC -!- CATALYTIC ACTIVITY: UDP-glucose + D-fructose 6-phosphate = UDP + sucrose 6-phosphate.

CC -!- ENZYME REGULATION: ACTIVITY REGULATED BY PHOSPHORYLATION AND MODERATED BY CONCENTRATION OF METABOLITES AND LIGHT.

CC -!- PATHWAY: Sucrose synthesis.

CC -!- SUBUNIT: HOMODIMER OR HOMOTETRAMER (BY SIMILARITY).

CC -!- PTM: PHOSPHORYLATED. HOWEVER, PHOSPHORYLATION IS NOT ESSENTIAL FOR ENZYME FUNCTION (BY SIMILARITY).

CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 1.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: U33175; AAC49379.1; -

DR InterPro: IPR001296; Glycos\_transf\_1.

DR Pfam: PF00534; Glycos\_transf\_1; 1.

KW Transferase; Glycosyltransferase; Phosphorylation.

FT DOMAIN 22 29 POLY-GLY.

FT DOMAIN 695 698 POLY-GLU.

FT DOMAIN 775 779 POLY-ARG.

SQ SEQUENCE 1049 AA; 116455 MW; ED862E2819AA4B04 CRC64;

Query Match 50.0%; Score 37; DB 1; Length 1049;

Best Local Similarity 57.1%; Pred. No. 65;

Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 ARRGTKCLREAPR 14

Db 420 ARRGVSGHGRFMR 433

||||| | | | |

## RESULT 15

SPS2\_CRAPL STANDARD; PRT; 1081 AA.

AC Q04933;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Sucrose-phosphate synthase 2 (EC 2.4.1.14) (UDP-glucose-fructose-phosphate glucosyltransferase 2).

DE SPS2.

GN

OS Craterostigma plantaginum.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Asteridae; euasterids I; Lamiales; Lamiales incertae sedis; Toreneae;

OC Craterostigma.

OX NCBI\_TaxID=4153;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=97451773; PubMed=9306694;

RA Ingram J., Chandler J.W., Gallagher L., Salami F., Bartels D.;

RT "Analysis of cDNA clones encoding sucrose-phosphate synthase in relation to sugar interconversions associated with dehydration in the resurrection plant Craterostigma plantaginum Hochst.";

RL Plant Physiol. 115:113-121(1997).

CC -!- FUNCTION: INVOLVED IN THE REGULATION OF CARBON PARTITIONING IN THE LEAVES OF PLANTS. MAY REGULATE THE SYNTHESIS OF SUCROSE AND THEREFORE PLAY A MAJOR ROLE AS A LIMITING FACTOR IN THE EXPORT OF PHOTOASSIMILATES OUT OF THE LEAF.

CC -!- CATALYTIC ACTIVITY: UDP-glucose + D-fructose 6-phosphate = UDP + sucrose 6-phosphate.

CC -!- ENZYME REGULATION: ACTIVITY REGULATED BY PHOSPHORYLATION AND MODERATED BY CONCENTRATION OF METABOLITES AND LIGHT.

CC -!- PATHWAY: Sucrose synthesis.

CC -!- SUBUNIT: HOMODIMER OR HOMOTETRAMER (BY SIMILARITY).

CC -!- PTM: PHOSPHORYLATED. HOWEVER, PHOSPHORYLATION IS NOT ESSENTIAL FOR ENZYME FUNCTION (BY SIMILARITY).

CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 1.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: Y11795; CAA72491.1; -

DR InterPro: IPR001296; Glycos\_transf\_1.

DR Pfam: PF00534; Glycos\_transf\_1; 1.

KW Transferase; Glycosyltransferase; Phosphorylation; Multigene family.

FT DOMAIN 245 248 POLY-SER.

FT DOMAIN 256 264 POLY-GLU.

FT DOMAIN 787 790 POLY-ARG.

SQ SEQUENCE 1081 AA; 120933 MW; DDL42DC2F1A72900 CRC64;

Query Match 50.0%; Score 37; DB 1; Length 1081;

Best Local Similarity 57.1%; Pred. No. 67;

Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 ARRGTKCLREAPR 14

Db 429 ARRGVNHGRFMR 442

||||| | | | |

Search completed: July 8, 2003, 16:25:06

Job time : 7.82051 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model  
Run on: July 8, 2003, 16:17:10 ; Search time 30.1538 Seconds  
(without alignments)  
95.665 Million cell updates/sec

Title: US-09-824-647-7  
Perfect score: 74  
Sequence: 1 ARRGTKCLREAPR 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL.21.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_virus.\*  
16: sp\_bacteriopl.\*  
17: sp\_archaea.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	68	91.9	413	Q9H8S1	Q9H8S1 homo sapien
2	48	64.9	589	11 Q9D2V3	Q9D2V3 mus musculus
3	44.5	60.1	694	16 Q8VJW0	Q8VJW0 mycobacteri
4	43	58.1	78	10 Q9XGJ0	Q9XGJ0 hordeum vul
5	42.5	57.4	546	11 Q9C9Q1	Q9C9Q1 mus musculus
6	42.5	57.4	546	11 Q9D2P9	Q9D2P9 mus musculus
7	41	55.4	130	11 Q9WUQ4	Q9WUQ4 rattus norv
8	41	55.4	429	4 O43729	O43729 homo sapien
9	41	55.4	433	11 Q9D603	Q9D603 mus musculus
10	41	55.4	1093	10 Q9SN30	Q9SN30 arabidopsis
11	40	54.1	327	5 Q93345	Q93345 caenorhabdi
12	40	54.1	422	2 Q9ZG90	Q9ZG90 flavobacter
13	40	54.1	500	2 Q9JPD5	Q9JPD5 rhodocyclu
14	40	54.1	831	17 Q9HNA5	Q9HNA5 halobacteri
15	39	52.7	100	13 Q8QGB6	Q8QGB6 oncorhynch
16	39	52.7	118	4 Q9Y3V0	Q9Y3V0 homo sapien

17	39	52.7	190	2	O31034	O31034 mycobacteri
18	39	52.7	265	11	Q9DCP5	Q9DCP5 mus musculus
19	39	52.7	278	11	Q9QXT5	Q9QXT5 mus musculus
20	39	52.7	395	10	Q947T9	Q947T9 oryza sativ
21	39	52.7	452	2	Q8VVA1	Q8VVA1 burkholderi
22	39	52.7	464	4	Q8TCW8	Q8TCW8 homo sapien
23	39	52.7	548	4	Q8WV46	Q8WV46 homo sapien
24	38	51.4	187	16	Q9RY33	Q9RY33 deinococcus
25	38	51.4	206	11	Q62912	Q62912 rattus norv
26	38	51.4	300	12	P89468	P89468 herpes simp
27	38	51.4	409	12	Q9DYB6	Q9DYB6 pseudorabie
28	38	51.4	465	10	Q8W052	Q8W052 oryza sativ
29	37	50.0	50	6	Q9GK00	Q9GK00 ateles sp.
30	37	50.0	99	12	Q91FT0	Q91FT0 chilo iride
31	37	50.0	114	4	O95014	O95014 homo sapien
32	37	50.0	134	10	Q9C703	Q9C703 arabidopsis
33	37	50.0	157	10	O49178	O49178 oryza sativ
34	37	50.0	203	11	Q9D9M1	Q9D9M1 mus musculus
35	37	50.0	246	16	Q8VT01	Q8VT01 anabaena sp
36	37	50.0	288	5	Q22385	Q22385 caenorhabdi
37	37	50.0	305	5	Q8T8W0	Q8T8W0 drosophila
38	37	50.0	316	4	Q9BS90	Q9BS90 homo sapien
39	37	50.0	353	11	Q9CZE0	Q9CZE0 mus musculus
40	37	50.0	372	5	Q9V4S3	Q9V4S3 drosophila
41	37	50.0	376	16	Q9RYP2	Q9RYP2 deinococcus
42	37	50.0	379	16	Q8YMU7	Q8YMU7 anabaena sp
43	37	50.0	385	4	Q9NXX6	Q9NXX6 homo sapien
44	37	50.0	400	11	Q8VDS7	Q8VDS7 mus musculus
45	37	50.0	402	11	Q9JTB2	Q9JTB2 rattus norv

ALIGNMENTS

RESULT 1

Q9H8S1 PRELIMINARY; PRG; 413 AA.  
AC Q9H8S1; 2001 (TREMREL. 16, Created)  
DT 01-MAR-2001 (TREMREL. 16, Last sequence update)  
DT 01-JUN-2002 (TREMREL. 21, Last annotation update)  
DE OVARC1001154 protein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=OVARIAN CARCINOMA;  
RA Isoqal T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
RA Nishikawa T., Nagai K., Sato H., Sugano S., Shiratori A., Sudo H.,  
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,  
RA Ninomiya K., Iwamoto T.,  
RA "NEDO human cDNA sequencing project."  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK023348; BAB14535.1;  
DR InterPro; IPR000118; Granulin.  
DR InterPro; IPR001211; PhospholipaseA2.  
DR Pfam; PF00396; granulin; 4.  
DR SMART; SM00277; GRAN; 4.  
DR PROSITE; PS00799; GRANULINS; UNKNOWN\_4.  
DR PROSITE; PS00118; PA2\_HIS; UNKNOWN\_1.  
SQ SEQUENCE 413 AA; 44132 MW; 0E3767A44BE314EC CRC64;

Query Match 91.9%; Score 68; DB 4; Length 413;  
Best Local Similarity 92.9%; Pred. NO. 0.00024;  
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ARRGTKCLREAPR 14  
| | | | | | | | | | | | | | | |

DB 386 RRGTKCLRRREAPR 399

RESULT 2

Q9D2V3 PRELIMINARY; PRT; 589 AA.

AC Q9D2V3

DT 01-JUN-2001 (TREMELrel. 17, Created)

DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)

DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)

DE Adult male kidney cDNA, RIKEN full-length enriched library,

DE clone:0610012H06, full insert sequence.

GN GRN.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=KIDNEY;

RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,

RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavani T.,

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

RA Krahl L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

RA Nordone P., Ring B., Rignwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schoenbach C., Seta Y., Shibata Y., Storch K.-F.,

RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,

RA Hayashizaki Y., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,

RA Rayshaw-Boris A.;

RT Functional annotation of a full-length mouse cDNA collection.\*;

RL Nature 409:685-690(2001).

DR EMBL; AK018744; BAB31384.1; -

DR MGD; MGI:95832; Grn.

DR InterPro: IPR000118; Granulin.

DR InterPro: IPR001211; PhospholipaseA2.

DR Pfam; PF00396; granulin; 7.

DR SMART; SM00277; GRAN; 7.

DR PROSITE; PS00799; GRANULINS; 7.

DR PROSITE; PS00118; PA2\_HIS; UNKNOWN.1.

SQ SEQUENCE 589 AA; 63405 MW; 1DE8229CA13CA292 CRC64;

Query Match 64.9%; Score 48; DB 11; Length 589;

Best Local Similarity 72.7%; Pred. No. 1.5;

Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RGTCLRRREAPR 13

DB 564 RGTCLRRKIP 574

RESULT 3

Q8VJW0 PRELIMINARY; PRT; 694 AA.

AC Q8VJW0

DT 01-MAR-2002 (TREMELrel. 20, Created)

DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)

DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)

DE PPE family protein.

GN Wt1857.

OS Mycobacterium tuberculosis.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.

OX NCBI\_TaxID=1773;

RN SEQUENCE FROM N.A.

RP STRAIN=CDC 1551 / OSHKOSH;

RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,

RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,

RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,

RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,

RA Bishai W.;

RT "Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains.";

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AS007044; AAK46130.1; -

DR TIGR; Wt1857; -

DR InterPro: IPR000568; ATPsint\_Asub.

DR InterPro: IPR000030; Microbac\_PPE.

DR Pfam; PF00823; PPE; 1.

DR PROSITE; PS00449; ATPASE\_A; UNKNOWN.1.

SQ SEQUENCE 694 AA; 73206 MW; 2FDAD94290A07359 CRC64;

Query Match 60.1%; Score 44.5; DB 16; Length 694;

Best Local Similarity 69.2%; Pred. No. 7.7;

Matches 9; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 2 RRGTKCLRRREAPR 13

DB 126 RRGTKCLRRRDPG 138

RESULT 4

Q9XGJ0 PRELIMINARY; PRT; 78 AA.

AC Q9XGJ0

DT 01-NOV-1999 (TREMELrel. 12, Created)

DT 01-NOV-1999 (TREMELrel. 12, Last sequence update)

DT 01-NOV-1999 (TREMELrel. 12, Last annotation update)

DE Sucrose-phosphate synthase 1 (Fragment).

GN SP81.

OS Hordeum vulgare (Barley).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;

OC Triticeae; Hordeum.

OX NCBI\_TaxID=4513;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. BARAKA; TISSUE=PRIMARY LEAF BLADES;

RA Mueller J., Aeschbacher R.A., Sprenger N., Boeller T., Wlémken A.;

RT "Carbohydrate-mediated regulation of two key enzymes of fructan synthesis in barley (Hordeum vulgare L.cv. Baraka) leaves.";

RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ238715; CAB45560.1; -

FT NON\_TER

FT NON\_TER

SQ SEQUENCE 78 AA; 8821 MW; BC4BF1DB9BE10AD0 CRC64;

Query Match 58.1%; Score 43; DB 10; Length 78;

Best Local Similarity 61.5%; Pred. No. 1.8;

Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 RRGTKCLRRREAPR 14

DB 1 RRGVSCYGREMPR 13

RESULT 5

Q9CY01 PRELIMINARY; PRT; 546 AA.

AC Q9CY01

DT 01-JUN-2001 (TREMELrel. 17, Created)

DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)

DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)

DE 3930401K13RIK protein.

GN 3930401K13RIK.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=EMBRYO;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 DR EMBL; AK014456; BAB29363.1;  
 DR MGD; MGI:1921272; 393040IK13Rik.  
 DR InterPro; IPR001744; 6PGD.  
 DR InterPro; IPR000637; AT.hook.  
 DR InterPro; IPR000313; PWWP\_domain.  
 DR Pfam; PF02178; AT.hook; 1.  
 DR Pfam; PF03446; NAD\_binding\_2; 1.  
 DR Pfam; PF00855; PWWP; 1.  
 DR SMART; SM00384; AT.hook; 1.  
 DR SMART; SM00293; PWWP; 1.  
 SQ SEQUENCE 546 AA; 59744 MW; B3AC1562477ABCF1F CRC64;

Query Match 57.4%; Score 42.5; DB 11; Length 546;  
 Best Local Similarity 52.9%; Pred. No. 14;  
 Matches 9; Conservative 4; Mismatches 1; Indels 3; Gaps 1;  
 OY 1 ARRGTKCLRR---EAPR 14  
 DB 152 ADRGSKCLRAEQSPR 168  
 |||:||||:|::||  
 RESULT 6  
 ID Q922P9 PRELIMINARY; PRT; 546 AA.  
 AC Q922P9;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DE RIKEN cDNA 393040IK13 gene.  
 GN 393040IK13Rik.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Strausberg R.;  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC006893; AAH06893.1;  
 DR MGD; MGI:1921272; 393040IK13Rik.  
 DR InterPro; IPR001744; 6PGD.  
 DR InterPro; IPR000637; AT.hook.  
 DR InterPro; IPR000313; PWWP\_domain.  
 DR Pfam; PF02178; AT.hook; 1.  
 DR Pfam; PF03446; NAD\_binding\_2; 1.  
 DR Pfam; PF00855; PWWP; 1.

SQ SEQUENCE 546 AA; 59715 MW; F5D2090DE1F64723 CRC64;  
 Query Match 57.4%; Score 42.5; DB 11; Length 546;  
 Best Local Similarity 52.9%; Pred. No. 14;  
 Matches 9; Conservative 4; Mismatches 1; Indels 3; Gaps 1;  
 OY 1 ARRGTKCLRR---EAPR 14  
 DB 152 ADRGSKCLRAEQSPR 168  
 |||:||||:|::||  
 RESULT 7  
 ID Q9WUQ4 PRELIMINARY; PRT; 130 AA.  
 AC Q9WUQ4;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE Secretory leukocyte protease inhibitor.  
 GN SLPI.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SD;  
 RA Chen D.H., Xu X.P., Bagchi M.K., Bagchi I.C.;  
 RT "Molecular cloning and spatio-temporal expression of rat secretory  
 RT leukocyte protease inhibitor (SLPI) in the uterus.";  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF151982; AAD34035.1;  
 DR HSSP; P19957; 2REL.  
 DR InterPro; IPR002221; WAP.  
 DR Pfam; PF00095; Wap; 2.  
 DR PRINTS; PR00003; 4DISULPHCORE.  
 DR ProDom; PD001224; WAP; 1.  
 DR SMART; SM00217; WAP; 2.  
 DR PROSITE; PS00317; 4\_DISULFIDE\_CORE; 2.  
 KW Protease.  
 SQ SEQUENCE 130 AA; 14017 MW; A630BA3FCE3A9D9D CRC64;

Query Match 55.4%; Score 41; DB 11; Length 130;  
 Best Local Similarity 53.8%; Pred. No. 6.9;  
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 OY 1 ARRGTKCLRR---EAPR 13  
 DB 38 ARKPAQCLKREK 50  
 |||:||||:|::||  
 RESULT 8  
 ID Q43729 PRELIMINARY; PRT; 429 AA.  
 AC Q43729;  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE DJ388M5.4 (putative GS2 like protein).  
 GN DJ388M5.4.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Burgess J.;  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z97055; CAB09789.1;  
 DR InterPro; IPR002641; Patatin.  
 DR Pfam; PF01734; Patatin; 1.  
 SQ SEQUENCE 429 AA; 47912 MW; 6278FA8C7968BBC3 CRC64;



```

Query Match      55.4%; Score 41; DB 4; Length 429;
Best Local Similarity 70.0%; Pred. No. 21;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 5 TKCLREAPR 14
   |:|:|:|:|
Db 28 TCCLRRAPR 37

RESULT 9
Q9D603 PRELIMINARY; PRT; 433 AA.
AC Q9D603;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 4833426H19RIK protein.
GN 4833426H19RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=HEAD;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Haru A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Alizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo T., Niki S., Niki S., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK014771; BAB29543.1;
DR MGD: MGI:1923022; 4833426H19RIK.
DR InterPro: IPR002641; Patatin.
DR InterPro: IPR001230; Prenyl_site.
DR Pfam: PF01734; Patatin; 1.
DR PROSITE: PS00294; PRENYLATION; UNKNOWN 1.
SQ SEQUENCE 433 AA; 48423 MW; 05366EC8E129A0E CRC64;

Query Match      55.4%; Score 41; DB 11; Length 433;
Best Local Similarity 70.0%; Pred. No. 21;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 5 TKCLREAPR 14
   |:|:|:|:|
Db 28 TCCLRRAPR 37

RESULT 10
Q9SN30 PRELIMINARY; PRT; 1083 AA.
AC Q9SN30;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Sucrose-phosphate synthase-like protein (EC 2.4.1.14).
GN P28M11.40 OR AT4G10120.
OS Arabidopsis thaliana (Mouse-ear cress).

```

```

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Murphy G., Ridley P., Hudson S., Bancroft I., Meyers H.W.,
RA Mayer K.F.X., Lemcke K., Schueller C.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL049487; CAB39764.1;
DR EMBL: AL161516; CAB78135.1;
DR InterPro: IPR001296; Glycos_transf_1.
DR Pfam: PF00534; Glycos_transf_1; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 1083 AA; 122688 MW; EFC126FCA2137BB1 CRC64;

Query Match      55.4%; Score 41; DB 10; Length 1083;
Best Local Similarity 61.5%; Pred. No. 51;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 2 RRGTKCLREAPR 14
   |:|:|:|:|
Db 468 RRGVSCLGRTMPR 480

RESULT 11
Q93345 PRELIMINARY; PRT; 327 AA.
AC Q93345;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE C36B1.11 protein.
GN C36B1.11.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Lennard N.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL: Z80215; CAB02276.1;
SQ SEQUENCE 327 AA; 37139 MW; BF592D9F0A2FA022 CRC64;

Query Match      54.1%; Score 40; DB 5; Length 327;
Best Local Similarity 53.8%; Pred. No. 25;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 ARRGTKCLREAPR 13
   |:|:|:|:|
Db 274 SRRFENCIRRESP 286

RESULT 12
Q9ZG90 PRELIMINARY; PRT; 422 AA.
ID Q9ZG90
AC Q9ZG90;

```

DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Endo-beta-galactosidase.  
OS Flavobacterium keratolyticus.  
OC Bacteria; CF6 group; Flavobacteria; Flavobacteriaceae; Flavobacterium.  
OX NCBI\_TaxID=81421;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99051328; PubMed=9831650;  
RA Leng L., Zhu A., Zhang Z., Hurst R., Goldstein J.;  
RT "Cloning, functional expression and purification of endo-beta-  
galactosidase from Flavobacterium keratolyticus.";  
RL Gene 222:187-194(1998).  
DR EMBL: AF083896; RAD04036.1; -;  
DR InterPro: IPR003305; CBM\_CenC.  
DR InterPro: IPR000757; Glyco\_hydro\_16.  
DR Pfam: PF02018; CBM\_4.9; 1.  
DR Pfam: PF00722; Glyco\_hydro\_16; 1.  
SQ SEQUENCE 422 AA; 45783 MW; E072456A740D14AC CRC64;  
  
Query Match 54.1%; Score 40; DB 2; Length 422;  
Best Local Similarity 77.8%; Pred. No. 32;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 ARRGTKCLR 9  
Db 319 ARRGTKCIR 327  
  
RESULT 13  
ID Q9JPD5 PRELIMINARY; PRT; 500 AA.  
AC Q9JPD5;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE Exopolyphosphatase.  
GN ppx.  
OS Rhodococcus gelatinosus (Rhodopseudomonas gelatinosa).  
OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae;  
OC Rubrivivax.  
OX NCBI\_TaxID=28068;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-IL144;  
RA Nagashima K.V., Shimada K., Matsuura K.;  
RT "Phylogenetic analysis of photosynthetic genes of Rhodococcus  
gelatinosus: Possibility of horizontal gene transfer in purple  
bacteria.";  
RL Photosyn. Res. 36:185-191(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-IL144;  
RX MEDLINE=94132007; PubMed=8300574;  
RA Nagashima K.V., Matsuura K., Ohyama S., Shimada K.;  
RT "Primary structure and transcription of genes encoding B870 and  
photosynthetic reaction center apoproteins from Rubrivivax  
gelatinosus.";  
RL J. Biol. Chem. 269:2477-2484(1994).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-IL144;  
RA Igarashi N., Shimada K., Matsuura K., Nagashima K.V.;  
RT "Photosynthetic gene cluster in purple bacterium, Rubrivivax  
gelatinosus.";  
RL (In) Garab G. (eds.);  
RL Photosynthesis;  
RL mechanisms and effects (Proceedings of the 11th international congress  
on photosynthesis), pp.4:2889-2892, Kluwer Academic Publishers,  
Dordrecht (1999).  
RN [4]  
RP SEQUENCE FROM N.A.

RC STRAIN=IL144;  
RX MEDLINE=20031519; PubMed=10563807;  
RA Menin L., Yoshida M., Jaquinod M., Nagashima K.V., Matsuura K.,  
RA Parot P., Vermeglio A.;  
RT "dark aerobic growth conditions induce the synthesis of a high  
midpoint potential cytochrome c8 in the photosynthetic bacterium  
Rubrivivax gelatinosus.";  
RL Biochemistry 38:15238-15244(1999).  
DR EMBL: AB034704; BAA94021.1; -;  
DR InterPro: IPR003695; Ppx\_GppA.  
DR Pfam: PF02741; Ppx\_GppA; 1.  
SQ SEQUENCE 500 AA; 54839 MW; 4D301D723A243DEA CRC64;  
  
Query Match 54.1%; Score 40; DB 2; Length 500;  
Best Local Similarity 61.5%; Pred. No. 37;  
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
  
QY 2 RRGTKCLRREAPR 14  
Db 61 RRGTKCLRREAPR 73  
  
RESULT 14  
ID Q9HNA5 PRELIMINARY; PRT; 831 AA.  
AC Q9HNA5;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE MCM / cell division control protein 21.  
GN MCM OR VNG2181G.  
OS Halobacterium sp. (strain NRC-1).  
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;  
OC Halobacteriaceae; Halobacterium.  
OX NCBI\_TaxID=64091;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20504483; PubMed=11016950;  
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,  
RA Shukla H.D., Lasky S.R., Balliga N.S., Thorsson V., Shrogha J.,  
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,  
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,  
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,  
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,  
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,  
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;  
RT "Genome sequence of Halobacterium species NRC-1.";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).  
DR EMBL: AE005105; AAG20316.1; -;  
DR HSP; P72065; IAM2.  
DR InterPro: IPR003586; Hedgehog\_hntC.  
DR InterPro: IPR003587; Hedgehog\_hntN.  
DR InterPro: IPR002203; Intein.  
DR InterPro: IPR001208; MCM.  
DR Pfam: PF00493; MCM; 1.  
DR ProDom: PD001041; MCM; 2.  
DR SMART: SM00305; HintC; 1.  
DR SMART: SM00306; HintN; 1.  
DR SMART: SM00350; MCM; 1.  
DR PROSITE: PS00847; MCM\_1; UNKNOWN\_1.  
DR PROSITE: PS00051; MCM\_2; 2.  
DR PROSITE: PS00881; PROTEIN\_SPLICING; UNKNOWN\_1.  
KW Cell division; Complete proteome.  
SQ SEQUENCE 831 AA; 90593 MW; 84A805B091676970 CRC64;  
  
Query Match 54.1%; Score 40; DB 17; Length 831;  
Best Local Similarity 66.7%; Pred. No. 60;  
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 2 RRGTKCLRREAP 13  
Db 338 RRGTKCLRREAP 349

RESULT 15  
Q8QGB6 PRELIMINARY; PRT; 100 AA.  
AC Q8QGB6;  
DT 01-JUN-2002 (TReMBLrel. 21, Created)  
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
DE VHSV-induced protein-8.  
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
OX NCBI\_TaxID=8022;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Benmansour A., Boubinot P., Vaghefi N., O'Farrell C.;  
RT "Global Survey of Transcript Expression in Rainbow Trout Leukocytes  
RT Reveals a Major Contribution of Interferon Responsive Genes in the  
RT Early Response to a Rhabdovirus Infection."  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF483528; AM18467.1; -.  
SQ SEQUENCE 100 AA; 11203 MW; 8CB8F4FC779E5D1F CRC64;  
Query Match 52.7%; Score 39; DB 13; Length 100;  
Best Local Similarity 63.6%; Pred. No. 13;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 3 RGTKCLRREAP 13  
Db : ||| |||  
73 KGKKCLNPEAP 83

Search completed: July 8, 2003, 16:28:05  
Job time : 33.1538 secs